Helsinki University of Technology
Department of Electrical and Communications Engineering

Timo Peltonen

A Multichannel Measurement System for Room Acoustics Analysis

This Thesis has been submitted for official examination for the degree of Master of Science, in Espoo, on October 23rd, 2000.

Instructor of the Thesis

D.Sc.(Tech.) Tapio Lahti

Supervisor of the Thesis

Professor Matti Karjalainen

TEKNILLINEN KORKEAKOULU **DIPLOMITYÖN TIIVISTELMÄ**

Tekijä:	Timo Peltonen		
Työn nimi:	A Multichannel Measurement System for Room Acoustics Analysis		
Päivämäärä:	23.10.2000	Sivumäärä:	119
Osasto:	Sähkö- ja tietoliikennetekniikka		
Professuuri:	S-89 Akustiikka ja äänenkäsittelytekniikka		
Työn valvoja:	Professori Matti Karjalainen		
Työn ohjaaja:	TkT Tapio Lahti		

Tässä työssä käsitellään akustisten impulssivasteiden ja saliakustisten tunnuslukujen mittaukseen ja analysointiin liittyviä menetelmiä ja sovellutuksia.

Työssä esitellään lineaaristen ja aikainvarianttien (LTI-) järjestelmien teoriaa sekä impulssivasteen määrittämiseen soveltuvia menetelmiä. Lähemmin tarkastellaan MLS (maksimipituussekvenssi) -menetelmän ominaisuuksia ja sovellutuksia. Työssä perehdytään saliakustisten tunnuslukujen teoriaan ja mittausmenetelmiin. Lisäksi käsitellään todellisissa impulssivasteissa esiintyvien kohina- ja viivetekijöiden kompensointiin soveltuvia jälkikäsittelymenetelmiä.

Työn osana on suunniteltu ja toteutettu akustisten impulssivasteiden monikanavaiseen mittaukseen ja analysointiin soveltuva Matlab-pohjainen järjestelmä. Mittausjärjestelmän ratkaisuja ja ominaisuuksia kuvataan ja sen toimintaa arvioidaan eri tavoin testien ja vertailumittausten avulla.

Työssä pohditaan impulssivasteen mittaukseen, suodatukseen ja saliakustisten tunnuslukujen analyysiin liittyviä tekijöitä, keskittyen mittaus- ja analyysimenetelmien käytännön soveltuvuuteen ja rajoituksiin.

Avainsanat: maksimipituusjono, MLS, huoneakustiikka, saliakustiikka, akustinen impulssivaste, saliakustiset tunnusluvut, impulssivasteiden jälkikäsittely, akustinen mittaustekniikka

HELSINKI UNIVERSITY OF TECHNOLOGY **ABSTRACT OF THE MASTER'S THESIS**

Author:	Timo Peltonen
Name of the thesis:	A Multichannel Measurement System for Room Acoustics Analysis
Date:	23.10.2000 Number of pages: 119
Department:	Department of Electrical and Communications Engineering
Professorship:	S-89 Acoustics and Audio Signal Processing
Supervisor:	Professor Matti Karjalainen
Instructor:	D.Sc.(Tech.) Tapio Lahti

This thesis is focused on the methods and applications related to the measurement and analysis of acoustical impulse responses and room acoustical criteria.

A background review is made on LTI (linear time-invariant) systems and impulse response acquisition methods. An emphasis is laid on the properties and application of the MLS (maximum length sequence) method. Room acoustical criteria are discussed, along with a variety of measurement methods suitable for their determination. Processing methods are studied for the compensation of noise and delay artifacts found in real impulse responses.

A Matlab-based system for the multichannel measurement and analysis of room acoustical impulse responses has been designed and implemented as a part of this work. The design and properties of the system software and hardware are described. The measurement system is evaluated in several aspects with tests and comparative measurements.

Aspects of impulse response measurements, filtering and room acoustical analysis are discussed, with an emphasis on their practical applications and limitations.

Keywords: maximum length sequence, MLS, room acoustics, acoustic impulse response, room acoustic parameters, acoustical measurements

Preface

This work is a Master's Thesis carried out for the Laboratory of Acoustics and Audio Signal Processing of the Helsinki University of Technology (HUT). It has been carried out as a part of the TAKU/BINA† project of the VÄRE technology program, and has been funded by TEKES. The main collaborators for the TAKU/BINA project are HUT, Insinööritoimisto Akukon and Nokia.

The work has been instructed by D.Sc.(Tech.) Tapio Lahti. During the course of the work, he has provided numerous important views and experience in the many discussions we have taken on the subject. The thorough feedback given on various proof versions of this book has been invaluable to the author.

Mr. Henrik Möller has provided a wealth of room acoustical insight, and has acted as an important source of inspiration. Tapio Lokki, Benoit Goutarbes and Lauri Savioja have all participated with their time and advice while testing and experimenting with various stages of the IRMA measurement system.

I would like to thank my supervisor Professor Matti Karjalainen, instructor Tapio Lahti, Henrik Möller, and all the aforementioned people for all their efforts and cooperation at the various stages of this project.

I am also indebted to my family, who have given their abiding affection to me as a father only vaguely present for a prolonged period of time. Nevertheless, my children Sakari and Unna have already both shown a keen interest in reverberant spaces at a young age.

The financial support of HUT, Akukon and all the TEKES project participants is gratefully acknowledged.

Espoo, 23.10.2000

Timo Peltonen

† TAKU/BINA is an acronym for tila-akustiikka ja binauraalinen mallinnus, *i.e.* room acoustics and binaural modeling.

Table of Contents

List of symbols

A	absorption area
c	sound velocity
C	clarity
C_S	stage clarity
D	definition
f_s	sample rate
G	strength
$h(t)$, $h(n)$, $h'(n)$	impulse response
$H(f)$	frequency response function
$\boldsymbol{H}_n$	Hadamard matrix
$\boldsymbol{I}_n$	identity matrix
n	MLS sequence order, sample number index
L	sequence length
L_f	lateral energy fraction
LE	lateral efficiency
p	sound pressure
$\boldsymbol{P}$	permutation matrices
r	correlation coefficient
$R(t)$	reverberation decay
ST	stage support factors
T	time between successive samples
T_{10}	early decay time
T_{60}	reverberation time
T_{rise}	rise time
T_S	center time
V	room volume
$x(t)$, $x(n)$, $x'(n)$, $X(f)$	stimulus
$y(t)$, $y(n)$, $y'(n)$, $Y(f)$	response
$z'(n)$	periodic MLS
$R_{xx}(t)$	autocorrelation function
$R_{xy}(t)$	cross-correlation function
$\boldsymbol{W}$	Walsh-Hadamard matrix
$\delta(t)$	Dirac delta function
$\delta(n)$, $\delta'(n)$	unit-sample sequence
$\Theta(f)$	phase response
Φ	interaural cross-correlation function
$*$	convolution
$\otimes$	periodic convolution

List of abbreviations

A/D	analog-to-digital
ASW	apparent source width
BR	bass ratio
C	clarity
D	definition, deutlichkeit
D/A	digital-to-analog
DFT	discrete Fourier transform
DUT	device under test
EDT	early decay time
FFT	fast Fourier transform
FHT	fast Hadamard transform
IACC	inter-aural cross-correlation coefficient
IACF	inter-aural cross-correlation function
IFFT	inverse fast Fourier transform
IIR	infinite impulse response
IR	impulse response
IRMA	Impulse Response Measurement Application
LE	lateral efficiency
LEF	lateral energy fraction
LEV	listener envelopment
LTI	linear and time-invariant
MLS	maximum length sequence
PIR	periodic impulse response
PN, PNS	pseudo-noise, pseudo-noise sequence
PS	power spectrum
RASTI	rapid speech transfer index
RMS	root-mean-square
T	reverberation time
ST	stage support factor
STI	speech transfer index
TDS	time-delay spectrometry
XOR	exclusive-OR function

1 Introduction

The topic of this thesis is the measurement and analysis of impulse responses in acoustical systems.

The study of room acoustics can be categorized to empirical observations and their underlying physical theory. Acoustical measurements form the basis for empirically determining the acoustical qualities of rooms, halls and other spaces. Together with room acoustical theory, empirically acquired data forms the basis for studying, modeling and designing acoustical spaces for a wealth of applications.

There exist a large number of room acoustical parameters, which are commonly used in attempt to numerically describe various acoustical properties of rooms and halls. Most of these indices are related to subjective aspects, which provide methods for empirically determining and comparing what room spaces sound like.

As room acoustical systems generally exhibit extensively complex behavior, comprehensive measurement and analysis methods are required to provide an understanding of the various physical phenomena and the forms of their mutual interaction in these systems.

Rapidly evolving computer and audio technology provides prospects for novel approaches in the methods used in room acoustical measurements. Multichannel methods seem a promising alternative to traditional monaural room acoustical measurements and parameters.

The goal of this thesis is to develop a multichannel impulse response measurement system suitable for a wide range of applications in acoustical measurements and analysis.

1.1 Background

Applications of acoustical impulse response measurements

By definition, an acoustical impulse response describes exclusively the behavior of a linear and time-invariant (LTI) acoustical system between two points in space. The determination of standard room acoustical parameters can be formulated from the impulse response.

Most room acoustical parameters, such as reverberation time and energy-time relations, have been defined only for a response measured using a single sound source and a microphone, both of which exhibit omnidirectional properties. Due to local variations in the sound field, the parameters are in practice averaged over several responses measured at various locations, in order to obtain mean values describing the room in general.

The directional information of a sound field is commonly quantified by two-channel parameters such as the LEF (Lateral Energy Fraction) and the IACC (Inter-Aural Cross-Correlation). These parameters require two responses, ac-

quired either by using two microphones with different directivities, or by two responses acquired at different receiver locations.

The developments in the research on room acoustics and the growing interest in virtual acoustics have generated an increasing need for multichannel measurement and analysis methods. Multichannel impulse response measurements yield possibilities for a diverse range of acoustical measurements. Examples of these are binaural measurements, sound intensity measurements and a range of multidimensional sound field measurements applying microphone rows, matrices, or three-dimensional rigs. The possibility of simultaneously acquiring multiple responses is also of great practical importance, as it effectively decreases the time required for extensive multichannel measurements. This introduces a whole new category of measurements, which would be practically unrealizable using a single-channel measurement system.

Multichannel room acoustic measurements offer new possibilities for the investigation and analysis of certain aurally distinguishable acoustical flaws, which traditional single channel methods and room acoustical parameters are not adapted to. For instance, accurate determination of the direction and source of individual reflections and echos in geometrically complex acoustic spaces could provide a valuable tool for practical room acoustical work.

1.2 Overview of the Thesis

Scope

The scope of this work has been to design and build a multichannel system for the measurement and analysis of acoustical impulse responses, and to accomplish an insight on the subject of current room acoustical measurements and the analysis methods involved. The study of background theory and a number of specialized measurement and analysis methods, together with the development, realization and practical application of the measurement system software and hardware, have caused this work to gain considerable extent.

Focus

The focus of this work has been to develop a portable multichannel system for acoustical impulse response measurement and analysis using standard studio sound hardware components and a Windows PC computer. An emphasis has been made on the adaptability and expandability of both the system hardware and software. By comparison, most commercial systems suited for impulse response measurements under field conditions are currently either limited to one or two channels, or exhibit a closed architecture.

Several key points have been concentrated on in the system integration: an open architecture in both the hardware and the software implementations, practical suitability for complex field measurements such as concert halls, and the possibility of expanding and adapting the system to a wide range of measurement applications. The system was built using ordinary computer audio and sound studio equipment, which effectively reduced the hardware costs as compared to the price of specialized and proprietary systems.

An emphasis is made on the study of processing methods for the correct analysis of results from real, unideal response data. This study has incorporated a close examination of available literature and publications related on the subject. Several processing methods described in these publications have been implemented in the analysis software.

The measurement system under discussion has been developed as a Master's Thesis project for the Laboratory of Acoustics and Audio Signal Processing at the Helsinki University of Technology. The work has been accomplished as a part of the VÄRE/TAKU project sponsored by TEKES.

The overall goal has been to provide a multichannel measurement and analysis system to be used in subsequent acoustical research work for the TEKES VÄRE technology program and its collaborators. The development of this system can be seen as the main contribution of the author.

Structural overview of the thesis

This thesis is structured as follows. An overview of impulse responses and their applications to room acoustics is given in Chapter 2. Chapter 3 deals with the methods applied for the measurement system. In Chapter 4, the results simulated and measured with the system are studied. Chapter 5 is devoted to the analysis of these results and a discussion of the realized system properties. Conclusions are drawn and directions for future work are given in Chapter 6.

Chapter 2 describes the theoretical basis for the methods and analysis contained in the rest of this work.

The theory of impulse responses and LTI (Linear and Time Invariant) systems is described. An emphasis is given to maximum length sequences and their properties, which form a powerful method for the accurate determination of impulse responses in LTI systems.

Impulse response analysis is covered with a description of standard room acoustic parameters, along with outlines on multichannel applications. Room acoustic parameters are used to describe the essential properties of acoustic spaces. These parameters are an attempt to distill acoustical impulse responses into single number indices describing a number of subjectively important acoustical properties.

In **Chapter 3**, the concepts and methods for devising a system for the measurement and analysis of acoustical impulse responses are discussed. A discussion on alternative measurement methods is included for comparison.

The methods and requirements for performing computer-based impulse response measurements are described. This is followed by a discussion of methods for processing, filtering and performing room acoustic analysis of impulse responses.

Chapter 4 describes the IRMA† measurement system, which has been designed and implemented by the author in the course of this study.

In **Chapter 5**, a series of measurements are performed in order to evaluate the IRMA system's performance and to verify its results.

The effects of normal and time-reversed filters on response decay are simulated, and the limits in the IIR filter realization are demonstrated.

The side effects of the maximum length sequence method are discussed. The effects of filtering, analysis and processing of responses for room acoustic analysis are shown with examples.

In **Chapter 6**, the results of room acoustic measurements performed in two acoustic spaces are presented and analyzed, providing comparative data between IRMA and the commercial MLSSA measurement system.

The IRMA measurement system is discussed in different aspects.

In **Chapter 7**, conclusions are given on the measurement and analysis methods applied, as well as on the design and realization of the measurement system and the results acquired.

† IRMA is an acronym for *Impulse Response Measurement Application*. The name relates to another feminine acronym, the MLSSA (pronounced Melissa) measurement system, which has been an important source of inspiration for this work.

2 Theory and background

This chapter lays the theoretical background on the use of impulse responses and maximum length sequences. With relations to acoustical decay theory, different room acoustical parameters are described, and applications in room acoustical and specialized measurements are given.

2.1 Impulse response theory

An *ideal physical system* is defined as a physically realizable system which is stable, linear, and has parameters of constant value [Bendat, Piersol 1980]. Such systems exhibit several important properties for linear analysis, which are described below.

2.1.1 Continuous-time systems

Figure 2.1 depicts a system producing an output $y(t)$ in response to an input $x(t)$. The output depends both on the input and the system response $h(t)$, which is a property of the system.

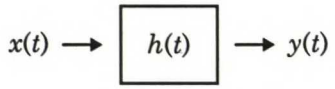


Figure 2.1. An ideal LTI system.

The *convolution integral* relates a system's output $y(t)$ to any arbitrary input $x(t)$ by the system's impulse response $h(t)$:

$$\begin{aligned} y(t) &= x(t) * h(t) \\ &= \int_0^\infty x(\tau) h(t-\tau) d\tau \end{aligned} \tag{2.1}$$

Causality limits the lower integration time limit to 0, as a *physically realizable system* cannot produce an output before an input is applied.

The requirement for *stability* limits the system's output to finite signals for all finite input signals.

The system is *time invariant*, if its parameters are constant in value. Thus the system response $h(t)$ is constant for all time values.

A *linear system* is defined as being additive and homogeneous. Let the system produce the output signals y_1 and y_2 for input signals x_1 and x_2 respectively. An *additive system* produces a summed output $(y_1 + y_2)$ for a summed input $(x_1 + x_2)$. A *homogeneous system* produces an output cy_1 for an input cx_1, where c is an arbitrary constant.

A combination of the last two properties is referred to as a *linear and time invariant (LTI)* system. These properties will be of major importance in most of the discussion that follows.

The impulse response

The *unit impulse response function* $h(t)$ is defined as a system's response to an ideal impulse, i.e. the *Dirac delta function* $\delta(t)$ [Bendat, Piersol 1980]. The unit impulse response function fully describes the system's behavior for any response:

$$x(t)=\delta(t)\Rightarrow h(t)=y(t), \tag{2.2}$$

where time $t = 0$ at the moment when the delta function is input to the system. For brevity, $h(t)$ is hereby called the *impulse response (IR)*.

Autocorrelation and cross-correlation

The autocorrelation function of the input signal is defined as

$$R_{xx}(t)=\int_0^\infty x(\tau)x(\tau+t)d\tau, \tag{2.3}$$

and the cross-correlation function of the input and output signals is expressed by

$$R_{xy}(t)=\int_0^\infty x(\tau)y(\tau+t)d\tau. \tag{2.4}$$

The convolution integral of equation 2.1 can be expanded to relate the autocorrelation and cross-correlation functions with each other in a similar fashion to $x(t)$ and $y(t)$ [Ando 1985]:

$$\begin{aligned} R_{xy}(t)&=\int_0^\infty h(\tau)R_{xx}(\tau-t)d\tau \\ &=h(t)*R_{xx}(t) \end{aligned} \tag{2.5}$$

For a signal with a constant power spectrum, such as white noise, the autocorrelation function is equal to a unit impulse:

$$R_{nn}(t)=\delta(t) \tag{2.6}$$

Application of such a signal as the system input $x(t)$ enables to obtain the system response $h(t)$ directly:

$$\begin{aligned} R_{xy}(t)&=\int_0^\infty h(\tau)\delta(\tau-t)d\tau \\ &=h(t)*\delta(t) \\ &=h(t) \end{aligned} \tag{2.7}$$

Any linear and time invariant system may be fully described by its impulse response $h(t)$. This relationship can naturally be stated in the frequency domain as complex division, which forms a practical method of calculating the system response [Bendat, Piersol 1980]:

$$\begin{aligned} \boldsymbol{Y}(f)&=\boldsymbol{H}(f)\boldsymbol{X}(f) \\ \boldsymbol{H}(f)&=\frac{\boldsymbol{Y}(f)}{\boldsymbol{X}(f)} \end{aligned}, \tag{2.8}$$

where $X(f)$, $Y(f)$ and $H(f)$ are the Fourier transforms of $x(t)$, $y(t)$ and $h(t)$ correspondingly.

2.1.2 Discrete-time systems

The definitions above apply for continuous-time systems only. The same properties may be defined for the analysis of discrete-time sampled data [Oppenheim, Schafer 1975].

The discrete-time version of the Dirac delta function is the *unit-sample sequence* defined as

$$\delta(n)=\begin{cases}1, n=0\\0, n\neq 0\end{cases} \tag{2.9}$$

The output sequence $y(n)$ of an LTI discrete-time system is related to an input sequence $x(n)$ by the system's unit-sample response $h(n)$ via *discrete linear convolution*,

$$\begin{aligned}y(n)&=x(n)*h(n)\\&=\sum_{k=-\infty}^{\infty}x(k)h(n-k)\end{aligned} \tag{2.10}$$

Using the unit sample sequence $\delta(n)$ as the stimulus $x(n)$, the system response $y(n)$ is equal to the unit-sample response $h(n)$, which fully describes the system. Note that for convenience, the unit-sample response is called the impulse response in the following discrete-time context. Although this nomenclature is mathematically inaccurate, the exact meaning can always be inferred from the context.

2.1.3 Periodic discrete-time systems

For a *periodic sequence* of length L, the *periodic unit-sample sequence* is [Rife, Vanderkooy 1989]

$$\delta'(n)=\begin{cases}1, n \bmod L=0\\0, n \bmod L\neq 0\end{cases} \tag{2.11}$$

Periodic sequences are hereby marked with an apostrophe to distinguish them from related aperiodic sequences.

Using *periodic convolution*, an LTI system can be characterized accordingly as follows:

$$\begin{aligned}y'(n)&=x'(n)\otimes h'(n)\\&=\sum_{k=0}^{L-1}x'(k)h'(n-k)\end{aligned} \tag{2.12}$$

Here $y'(n)$, $x'(n)$, and $h'(n)$ are all periodic sequences with period L, and the indices are all evaluated modulo L. Henceforth all indices of periodic sequences $z'(n)$ are evaluated modulo L, unless otherwise noted. The symbol $\otimes$ is used for periodic convolution to distinguish it from regular convolution, which is marked with the $*$ symbol.

$h'(n)$ is referred to as the *periodic impulse response (PIR)*.

The relationship between the impulse response $h(n)$ and the periodic impulse response $h'(n)$ can be determined as follows:

$$\begin{aligned} h'(n) &= \sum_{k=-\infty}^{\infty} \delta'(k) h(n-k) \\ &= \sum_{k=-\infty}^{\infty} h(n+kL) \end{aligned} \tag{2.13}$$

This convolution of $\delta'(n)$ with $h(n)$ effectively wraps L-point segments of the nonperiodic impulse response $h(n)$ to the origin, and accumulates them to form the periodic impulse response $h'(n)$. If the segment length L is chosen such that the impulse response $h(n)$ decays to a negligible value over the first L samples, time aliasing is avoided and the periodic impulse response $h'(n)$ closely approximates the first L samples of the actual impulse response. [Rife, Vanderkooy 1989]

If the nonperiodic response has significant content beyond the first L samples, the remainder is wrapped to the beginning of the periodic impulse response in L sample segments and accumulated together with the first L samples. This phenomenon is known as *time aliasing*, and should be avoided in ordinary response measurements. There are, however, special techniques such as interleaved high frequency oversampling [Mommertz, Bayer 1995], [Xiang, Genuit 1996], where time aliasing is effectively exploited.

It is important to realize that periodic methods for impulse response measurement always result in the periodic impulse response. It is thus essential to choose the sequence length L appropriately in order to avoid time aliasing. This requires a basic knowledge of the system decay time, or the use of sufficiently long sequences.

2.2 Maximum Length Sequences

The use of *maximum length sequences* (MLS) forms a powerful method for the accurate determination of impulse responses in LTI systems. The method is based on the use of a deterministic pseudo-random stimulus, which is cross-correlated with the acquired response to yield the impulse response of the system under test.

Maximum length sequences exhibit a number of favorable properties. The sequences have a minimal crest factor† , which results in theoretically optimal signal-to-noise ratios. The stimulus is deterministic, providing exact repeatability, and periodic, eliminating the need for time windowing. The MLS method is suitable for exceedingly long impulse responses, from which very finely scaled frequency responses can be calculated.

† The *crest factor* is defined as the the ratio of signal peak to RMS value [Carlson 1986].

2.2.1 General properties

Maximum length sequences are a type of pseudorandom sequences with a number of useful mathematical properties. In literature, maximum length sequences are also known as *pseudo-noise* (PN) *sequences*, *maximal-length shift-register sequences* or *m-sequences*. Although maximum length sequences with ternary or more states do exist [MacWilliams, Sloane 1976], the discussion here is limited to binary sequences.

Binary maximum length sequences have a period of $2^n - 1$, where n is a positive integer. The sequences are generated recursively using a digital shift register with n binary elements. Shift register output is produced at the last register element. The output and a number of other register elements called taps are summed together using the bitwise exclusive-or (XOR) function, and the result is fed back into the first element after shifting the register contents one element towards the output.

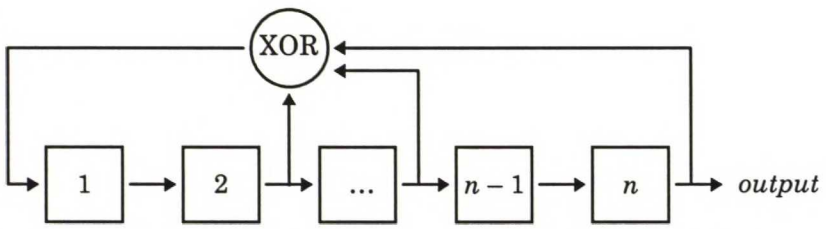


Figure 2.2. Shift register for generating maximum length sequences.

A binary shift register with n elements has 2^n discrete states. The shift register has two alternate modes of recursive operation: the null state, and the set of all the other register states. In the null state, all the register elements contain zeros. This can only lead to another null state, since 0 XOR 0 = 0. With the exception of the null state, the register cycles through the other $2^n - 1$ states recursively. None of the non-null states lead to the null state, so the register can be initialized in any of the non-null states for sequence generation. [MacWilliams, Sloane 1976]

The initial register state defines the starting point of the recursive sequence. It is customary to initialize all the register elements with 1's in order to define a static starting point for each sequence.

When concatenated together, the output bits of an n element shift register form a maximum length sequence of length $2^n - 1$. The number and locations of the feedback tap elements on the shift register determine the order of the $2^n - 1$ discrete register states, and thus changing the taps leads to different sequences of the same length. These sequences are mutually uncorrelated, which leads to potential applications in simultaneous multichannel measurements.

The number of 1's in a maximum length sequence exceeds the number of 0's exactly by one. Thus the sum of a binary (0,1) sequence is always $2^n - 1$.

Only certain shift register tap and order combinations cycle through all of the $2^n - 1$ states, and lead to maximum length sequences. Suitable tap combinations for common register lengths are available in literature ([MacWilliams, Sloane 1976], [Borish 1985], [Kovitz 1992], [Vanderkooy 1994]), some of which are

listed in Appendix A. Thorough discussions on the number theoretical properties of the MLS can be found in [MacWilliams, Sloane 1976] and [Lempel *et al.* 1977].

In practical applications, the binary MLS is commonly mapped from values (0, 1) to symmetrical signal levels (1, –1), which is called a *symmetrical MLS*. Due to the sequence properties described above, the sum of a symmetrical MLS is always –1. This results in an important factor for impulse response measurements: the symmetrical MLS signal is practically AC coupled, and contains very little energy at DC.

2.2.2 Discrete-time properties of the MLS

Maximum length sequences exhibit several important properties for applications in impulse response measurement:

The periodic autocorrelation of a symmetrical maximum length sequence $z'(n)$ with a period L is essentially a periodic unit-sample sequence:

$$R'_{zz}(n) = \frac{1}{L}\sum_{k=0}^{L-1} z'(k)z'(k+n) \tag{2.14}$$

$$\begin{aligned} R'_{zz}(0) &= 1 \\ R'_{zz}(n) &= -\frac{1}{L},\ 0 < n < L \end{aligned} \tag{2.15}$$

The MLS offers the optimal autocorrelation function of any binary sequences with the period L [MacWilliams, Sloane 1976]. The autocorrelation function also provides a good measure of the signal's degree of pseudorandomness.

For mathematical convenience, it is desirable to normalize the autocorrelation by $(L + 1)$ instead of the factor L, creating an approximation:

$$R''_{zz}(n) = \frac{1}{L+1}\sum_{k=0}^{L-1} z'(k)z'(k+n) \tag{2.16}$$

$$\begin{aligned} R''_{zz}(0) &= \frac{L}{L+1} \\ R''_{zz}(n) &= -\frac{1}{L+1},\ 0 < n < L \end{aligned} \tag{2.17}$$

This can also be expressed as the sum of the periodic unit-sample sequence $\delta'(n)$ and a small DC component:

$$R''_{zz}(n) = \delta'(n) - \frac{1}{L+1} \tag{2.18}$$

As the period L becomes large with long sequences, the DC component in the second term approaches zero, and the expression approaches the ideal aperiodic unit-sample sequence $\delta(n)$.

Applying an MLS $z'(n)$ to an LTI system with a periodic impulse response $h'(n)$, the system output $y'(n)$ is expressed as the periodic convolution

$$\begin{aligned} y'(n) &= z'(n) \otimes h'(n) \\ &= \sum_{k=0}^{L-1} z'(k) h'(n-k) \end{aligned} \tag{2.19}$$

The periodic impulse response $h'(n)$ may be recovered by cross-correlating the system output $y'(n)$ with the input $z'(n)$ using periodic cross-correlation. The cross-correlation is again normalized by $(L + 1)$:

$$R''_{zy}(n) = \frac{1}{L+1} \sum_{k=0}^{L-1} z'(k) y'(n+k) \tag{2.20}$$

Substituting equation 2.19 into 2.20 and manipulating further, results in the following equation. The full derivation involved is given in [Rife, Vanderkooy 1989].

$$R''_{zy}(n) = R''_{zz}(n) \otimes h'(n) . \tag{2.21}$$

This results in the fact that the periodic cross-correlation $R''_{zy}(n)$ of the output and the input sequences is equal to the periodic convolution of the MLS autocorrelation $R''_{zz}(n)$ with the system's periodic impulse response. This is in accordance with the result obtained for white noise in the continuous time domain. By combining this result with equation 2.18, we end up with the following equation:

$$\begin{aligned} R''_{zy}(n) &= h'(n) - \frac{1}{L+1} \sum_{k=0}^{L-1} h'(k) \\ &= h'(n) - \frac{1}{L} \sum_{k=0}^{L-1} h'(k) + \frac{1}{L(L+1)} \sum_{k=0}^{L-1} h'(k) \end{aligned} \tag{2.22}$$

The second term in this equation is the mean value of the periodic impulse response, and represents the DC component. The last term is the same component scaled down by $1 / (L + 1)$. The second term removes the DC component from the periodic impulse response, leaving just the last term. This is also negligible for large L. If the measurement system is AC coupled, the sum of $h'(n)$ over one period is zero, and the second and third terms are cancelled out altogether. Any extraneous DC offsets in the measuring chain are also attenuated by a factor of $(L + 1)$.

Due to this AC coupling, results obtained using unit sample sequences and symmetrical MLS are not necessarily equal. It must be kept in mind that unlike unit sample sequences, symmetrical MLS measurements are virtually AC coupled, with no signal energy in the DC component. Fortunately this behavior has very little effect on the analysis of acoustical systems. For applications with DC coupled systems, a constant offset value of –1 may be added to the MLS stimulus, enabling complete recovery of the impulse response including the DC component [Rife, Vanderkooy 1989].

From a frequency-domain perspective, the MLS stimulus is found to exhibit a constant flat spectrum except at DC. The following equation shows a comparison of the power spectra of a unit sample sequence and the MLS. The *power spectrum (*PS*)* of a periodic sequence is defined as the *discrete Fourier transform (*DFT*)* of its autocorrelation [Rife, Vanderkooy 1989].

$$\mathrm{PS}(\delta'(n)) = \mathrm{DFT}(R_{\delta\delta}) = \frac{1}{L}$$
$$\mathrm{PS}(z'(n)) = \mathrm{DFT}(R''_{zz}) = \begin{cases} \frac{1}{L} & , n = 0 \\ \frac{L+1}{L} & , n \neq 0 \end{cases} \tag{2.23}$$

Ignoring the DC component, a symmetrical MLS is found to deliver $(L + 1) = 2^n$ times the signal power of a periodic unit-sample sequence with the same peak amplitude. This amplification accounts for the superior noise immunity in MLS measurements.

The phase spectra of the unit sample sequence and the MLS are given as

$$\Theta[\delta'(n)] = Arg[\mathrm{DFT}(R_{\delta\delta})] = 0$$
$$\Theta[z'(n)] = Arg[\mathrm{DFT}(R''_{zz})] \tag{2.24}$$

Unlike $\delta'(n)$, the phase spectrum of the MLS is not zero at all frequencies, but varies pseudo-randomly with frequency and has a uniform probability density over its range of $[-\pi, +\pi]$. [Rife, Vanderkooy 1989]

2.2.3 Properties of Hadamard matrices

Hadamard matrices and the Hadamard transform provide the basis for an efficient algorithm for calculating the periodic cross-correlation involved in MLS measurements [Borish, Angell 1983], [Lempel 1979].

Hadamard matrices are orthogonal $n \times n$ matrices with binary elements (–1, 1), with the following properties:

$$\boldsymbol{H}_n \boldsymbol{H}_n^T = \boldsymbol{H}_n^T \boldsymbol{H}_n = n \cdot \boldsymbol{I}_n, \tag{2.25}$$

where $\boldsymbol{I}_n$ is the identity matrix.

A Hadamard matrix multiplied by –1 is also a Hadamard matrix.

Permutating the rows and columns of a Hadamard matrix generates another Hadamard matrix. Any two Hadamard matrices $\boldsymbol{H}_A$ and $\boldsymbol{H}_B$ are considered equivalent, if

$$\boldsymbol{H}_B = \boldsymbol{P}_{rows}^T \boldsymbol{H}_A \boldsymbol{P}_{columns}, \tag{2.26}$$

where $\boldsymbol{P}_{rows}$ and $\boldsymbol{P}_{columns}$ are permutation matrices for the rows and columns of $\boldsymbol{H}_A$.

A special case of interest are Hadamard matrices of order 2^n, as there exists at least one equivalence class containing the different representations of the Walsh transform matrix, including a Walsh-Hadamard representation in natural Hadamard ordering. All other Walsh-Hadamard matrix representations of the same order can be obtained from the Hadamard ordering by permutation of the row and column numbers. [Sutter 1991]

A linear transformation by a Hadamard matrix is called the *Hadamard transform*. It is defined as the product of a Hadamard matrix and a vector. The *Walsh-Hadamard Transform*, which uses a Walsh-Hadamard matrix in natural ordering, can be rapidly computed by a simple *Fast Hadamard Transform (FHT)* algorithm [Gumas 1997]. This can be extended to a more general transform suitable for all Walsh-Hadamard matrix representations, by combining the Fast Walsh Transform with permutations to form the required representation.

Walsh-Hadamard matrices and the MLS have an important relation. A *Walsh-Hadamard matrix representation* of order 2^n may be constructed by filling a $(2^n - 1) \times (2^n - 1)$ matrix with the $2^n - 1$ cyclic shifts of an order n symmetrical MLS, and including a first row and column filled with 1's to form the resulting $2^n \times 2^n$ matrix. This follows from equation 2.25 and the autocorrelation property of the MLS (equation 2.15) [MacWilliams, Sloane 1976], [Lempel 1975].

This relation can be used to notably reduce the amount of computation required for the circular cross-correlation operation to calculate the periodic impulse response from an MLS measurement. The cross-correlation may be rewritten in matrix form

$$\boldsymbol{h} = \frac{1}{L+1} \boldsymbol{W} \cdot \boldsymbol{y} \tag{2.27}$$

where $\boldsymbol{y}$ is the periodic system response vector. $\boldsymbol{W}$ is a right-circulant matrix containing the n different states of the MLS, which can be represented by means of a Walsh-Hadamard matrix:

$$\boldsymbol{W} = \boldsymbol{P}_2 \boldsymbol{S}_2 \boldsymbol{H}_{L+1} \boldsymbol{S}_1 \boldsymbol{P}_1 , \tag{2.28}$$

where $\boldsymbol{H}_{L+1}$ is an order $L + 1$ Walsh-Hadamard matrix in natural ordering, and $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are permutation matrices used to obtain the Walsh-Hadamard matrix representation related to the MLS. Matrices $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are used to remove the first column and row of $\boldsymbol{H}$.

The efficiency of this operation is based on Fast Hadamard Transform algorithm, that can be computed in $O((L + 1) \log_2(L + 1))$ operations, whereas the direct cross-correlation requires an order of $O(L^2)$ operations [Kovitz, 1992]. An additional computational profit is that the whole transform reduces to operations of addition and subtraction, as the Hadamard matrix contains only unity values.

By combining the above two equations and grouping operations, the periodic impulse response can be calculated as follows:

$$\boldsymbol{h} = \frac{1}{L+1} \boldsymbol{P}_2 \Big(\boldsymbol{S}_2 \big(\boldsymbol{H}_{L+1} \big(\boldsymbol{S}_1 (\boldsymbol{P}_1 \cdot \boldsymbol{y}) \big) \big) \Big) \tag{2.29}$$

First, the system response vector $\boldsymbol{y}$ is reordered by permutation $\boldsymbol{P}_1$, and a zero element is affixed to the beginning of the vector. The resulting vector of length $L + 1$ is then transformed using the FHT algorithm [Gumas 1997]. The outcoming vector is truncated to length L by dropping the first element, followed by permutation $\boldsymbol{P}_2$. The resulting periodic impulse response $\boldsymbol{h}$ is obtained after normalization.

The permutations $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ share an important role in the calculation, enabling the conversion of the Walsh-Hadamard representation of a given MLS to the natural Walsh-Hadamard ordering required for the Fast Hadamard Transformation [Lempel 1979]. For a given MLS, the permutations need be calculated only once, commonly along with the generation of the sequence, and stored for later use with the FHT. The indices needed for the permutations can be calculated as follows.

The states of the MLS generating shift register are listed as matrix rows, and the values for $\boldsymbol{P}_1{}^T$ are formed by column using the elements in the first n rows in a binary fashion. The actual values used for $\boldsymbol{P}_1$ are found by calculating the transpose of $\boldsymbol{P}_1{}^T$.

The matrix is then permuted columnwise according to $\boldsymbol{P}_1{}^T$, so that the first n rows form an ascending index in a binary fashion. The values for $\boldsymbol{P}_2$ are then formed by row using the elements in columns $2^0 \ldots 2^n$ in a binary fashion. An example of this procedure is depicted in figure 2.3 for a sequence of order $n = 3$.

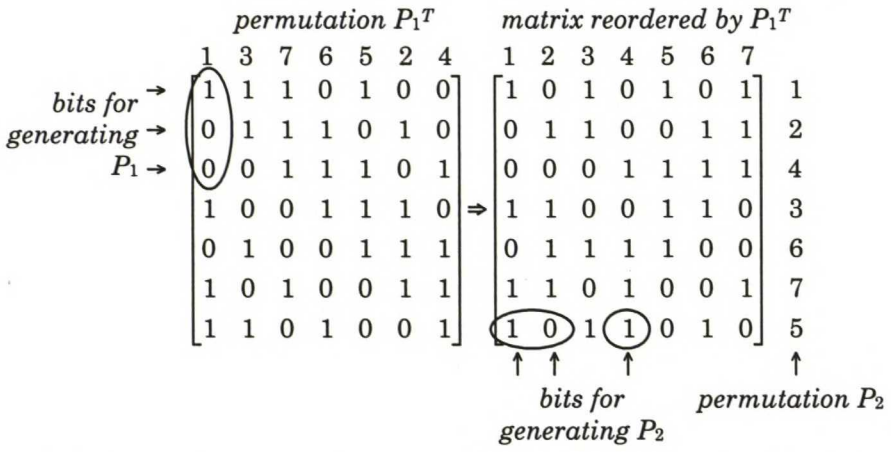


Figure 2.3. Determination of permutations P_1 and P_2. (After [Borish 1983]).

This method is somewhat impractical, considering the exceeding matrix sizes needed for sequences of higher orders. However, since the matrices only consist of cyclic shifts and permutations of the original MLS, it will suffice that single matrix elements are calculated directly from the MLS using modulo arithmetic for indexing. [Borish 1985], [Sutter 1991]

Equations for efficiently calculating $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are given below. All three equations use a range of $0 < i < L - 1$ for indexing. All vectors are indexed starting from zero. The corresponding maximum length sequence is denoted by $\boldsymbol{z}(i)$.

The transpose of permutation $\boldsymbol{P}_1$ is first calculated by

$$\boldsymbol{P}_1^T(i) = \sum_{j=0}^{n-1} 2^j \, z\left[(i + L - j)_{\bmod L}\right] \tag{2.30}$$

$\boldsymbol{P}_1{}^T$ is then transposed to give $\boldsymbol{P}_1$:

$$\boldsymbol{P}_1\left(\boldsymbol{P}_1^T(i) - 1\right) = i + 1. \tag{2.31}$$

Finally, the permutation $\boldsymbol{P}_2$ is calculated by

$$P_2(i) = \sum_{j=0}^{n-1} 2^j z\left[\left(P_1(2^j - 1) - i + L - 1\right)_{\bmod L}\right] \tag{2.32}$$

The above parameters enable the application of the FHT method for the calculation of periodic impulse responses according to equation 2.29.

2.2.4 Continuous-time properties

For analog measurements, the discrete-time MLS theory must be extended to continuous-time form. This is realized by mapping the discrete-time signals to continuous-time using a *unit-impulse comb function* convolved with a weighted *unit pulse function* $p(t)$, which is a train of unit samples spaced δt apart [Rife, Vanderkooy 1989]:

$$\mathrm{COMB}(t) * p(t) = \begin{cases} 1, & (n)\delta t < t < (n+1)\delta t \\ 0, & \text{otherwise} \end{cases} \tag{2.33}$$

In practice, this mapping is easily realized using a 1-bit sample-and-hold digital-to-analog converter, which holds its output constant over each sample period.

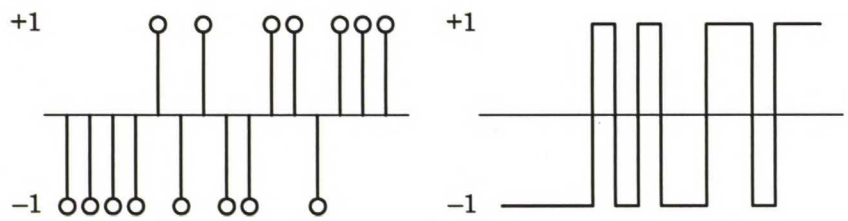


Figure 2.4a. Discrete-time MLS.
Figure 2.4b. Continuous-time MLS.

This mapping modifies the flat discrete-time MLS spectrum with the sinc(f)† spectrum of the unit pulse function $p(t)$‡. Following the D/A-conversion, the stimulus signal is lowpass filtered to prevent anti-aliasing.

For data acquisition, the analog system response must be sampled back to discrete-time data. This involves using an analog-to-digital converter, and another anti-aliasing lowpass filter, preferably similar to the first filter.

The resulting discrete-time periodic impulse response thus consists of the following elements:

$$y'(n) = \left\{\left[\left\{z'(n)\mathrm{COMB}(t)\right\} * p(t)\right] * h_{\text{system}}(t) * h_{\text{anti-alias}}(t)\right\}_{t=n\delta t} \tag{2.34}$$

---

† The sinc() function is defined as $\mathrm{sinc}(x) = \sin(x) / x$.

‡ In practice, this effect depends on the realization of the D/A converter. Many modern converters employ internal sinc compensation, resulting in a spectrally flat response together with a ringing behavior in the impulse response, as the sinc function is effectively transferred into the time domain.

The effects of the extraneous terms in the measured response may be somewhat compensated by first measuring a system loop-back response, comprising only of the extraneous elements, and deconvolving actual measured responses with it.

In the frequency domain, this comprises of dividing the complex system response by the complex loopback response. In practice, the main effect of compensating the response is the removal of the sinc(*f*) attenuation visible at high frequencies. With sinc compensated converters, no compensation is required in the frequency domain, but it can be used to remove the sinc(*f*) shaped ringing in the impulse response.

2.3 Room acoustic criteria

The theory of LTI systems, impulse responses, and the MLS method provide important tools for the measurement and description of acoustic spaces.

Acoustic impulse responses are temporal sound pressure functions of an acoustic space resulting from its excitation by an impulse that approximates the Dirac delta function. [Vorländer, Bietz 1994]

Impulse responses measured in acoustic spaces provide an accurate description of the acoustical properties of the system [Schroeder 1979]. The majority of room acoustic criteria, including all the standardized room acoustic parameters, can be calculated from acoustical impulse responses acquired using various measurement methods.

2.3.1 Measurement considerations

By definition, acoustical impulse responses are measured in source–receiver pairs. In room acoustical systems, the response between a single source and a receiver location fully describes the acoustic system between the two specific locations. It must be clarified though, that such results cannot be generalized to represent the complex acoustical system of the whole room space.

Spatial averaging

Due to local sound field variations, it is necessary to average data from multiple locally varied responses. Traditional measurement methods which apply moving sources, receivers or diffusor elements (suitable for reverberation chambers), are not applicable to impulse response measurements due to their time-variant nature. Impulse response measurements must thus be performed using a number of static source–receiver positions, placed at various locations around the room. The measurements between these pairs may be performed either consecutively or simultaneously.

Simultaneous measurements at multiple receiver points require the use of a multichannel acquisition system. Theoretically, it is even possible to measure responses from multiple simultaneous sources, if the applied individual stimuli are mutually uncorrelated. However, such a setup would result in rather adverse signal-to-noise conditions.

Directional issues

Standard room acoustic measurements are performed using omnidirectional transducers for the sources and receivers. This simplifies the definition, replication, measurement, and calibration of the transducers in order to acquire similar values with different pieces of measuring equipment.

The problem with omnidirectional measurements is their disregard to directional properties, which have important relations to the perceived acoustical qualities of spaces. Some of the newer room acoustical criteria, such as lateral energy and interaural cross-correlation [Ando 1985] are concentrated specifically on the directional properties of the response. These methods involve the acquisition of two locally and directionally varied responses.

Overview of room acoustical criteria

Acoustical spaces are commonly characterized by single-number indices that are calculated as an average of the responses acquired from the various source-receiver pairs.

In literature, these quantities are often referred to as *room acoustic parameters*. A name better suited for this context would be *room acoustic criteria* or *room acoustic properties*, which yield *room acoustic indices* as their values. This nomenclature will be used in this work.

Table 2.1 gives an overview of general room acoustic criteria that are either standardized or in common use.

Table 2.1. Overview of general room acoustic criteria.

quantity	symbol	subjective aspect	source
reverberation time (s)	T_{60}, T_{30}	reverberance	ISO 3382
early decay time (s)	T_{10}	reverberance, clarity	ISO 3382
center time (s)	T_S	clarity	ISO 3382
clarity (dB)	C	clarity	ISO 3382
definition (dB)	D	definition	ISO 3382
strength (dB)	G	relative level	ISO 3382
lateral energy fraction (%)	L_F, LEF	spatial impression	ISO 3382
interaural cross-correlation	IACC	spatial impression	ISO 3382

The definitions and applications of these quantities are discussed in the following sections.

2.3.2 Reverberation time

Reverberation time is defined as the time interval required for the sound energy level to decay 60 dB after excitation has stopped [Sabine 1900].

The ideal decay process

In a diffuse sound field, the ideal room decay process exhibits a purely exponential decay curve

$$p^2(t) = p_0^2 e^{-kt} \tag{2.35}$$

where p_0 is the sound pressure at zero time, and $p(t)$ is the sound pressure at time t. The decay parameter k is related to the room properties by

$$k = \frac{cA}{4V}, \tag{2.36}$$

where c = 340 m/s is the velocity of sound, A the total absorption area in the room, and V the volume of the room. According to Sabine, the *reverberation time* T_{60} is defined as

$$T_{60} = \frac{0.16V}{A}. \tag{2.37}$$

The coefficient 0.16 is empirically determined, and shows some variance with temperature. Other values such as 0.161 and 0.163 are also common in literature.

Combining equations 2.36 and 2.37, the parameter k is related to the reverberation time by

$$k = \frac{13.6}{T_{60}}. \tag{2.38}$$

The ideal room decay process may thus be rewritten as [Jordan 1980]

$$p^2(t) = p_0^2 e^{-\left(\frac{13.6t}{T_{60}}\right)}. \tag{2.39}$$

The reverberation time can be evaluated by plotting the decay curve on a logarithmic amplitude scale. Under ideal conditions, the decay curve forms a straight line, from which the time interval corresponding to the 60 dB of decay can be determined.

A complementary build-up process of the sound field may be written as

$$p^2(t) = p_0^2\left(1 - e^{-kt}\right) \tag{2.40}$$

This enables the definition of *rise time* T_{rise} corresponding to the point of time where 50% of the total energy has arrived [Jordan 1980]:

$$T_{\text{rise}} = \frac{\ln(0.5)}{k} T_{60} \approx 0.05 T_{60} \tag{2.41}$$

Factors affecting reverberation time measurement

A minimum source–receiver distance must be taken into account for room acoustic measurements. This is to ensure that the receiver is located in a reverberant field where the direct sound energy does not dominate in the response. The minimum distance can be calculated by [ISO 3382, 1997]

$$d_{\min} = 2\sqrt{\frac{V}{cT_{60}}}, \tag{2.42}$$

where V is the room volume, c is the speed of sound, and T_{60} is the approximated reverberation time.

For small spaces, this requirement poses problems. This is in accordance with the statistical diffuse-field definition of reverberation.

Practical considerations on decay curves

In practice, the properties of decay curves are not simply exponential. Background noise, natural room modes, and non-diffuse sound fields are the main factors for this behavior. The exponential decay is always obstructed by a constant level of background noise at some point. Natural room modes may cause the decay curve to warble, especially at low frequency bands where the mode density is low. In cases of an uneven distribution of absorption, the decay curve may bend from one slope to another, resulting in multiple levels of decay. This is more or less the case in almost all practical room spaces.

Since the calculation of reverberation time is based on an assumption of exponential decay, it is not necessary to acquire a full 60 dB of decay by measurement, but a smaller portion of the available dynamic range may be evaluated and the result simply scaled to 60 dB of decay. By using a normalized decibel scale on the amplitude axis, the evaluation and scaling can be performed simply by finding the slope of the decay curve, and extending it down to –60 dB.

Determination of the decay slope

Determination of the decay curve slope can be performed by a variety of methods. Traditionally the slopes were evaluated manually, as the data was only available on plotter paper. The evaluation could either be done by a visual "best fit" strategy†, or simply by finding the time difference between certain pre-determined levels on the curve, say from –5 dB to –35 dB. While the former method could be more accurate, the subjectiveness involved could cause deviations in judgement between observations of a single decay curve. The latter method is strictly repeatable, but it is inaccurate, if the decay curve exhibits warble or bends.

The ISO 3382 standard recommends that linear regression be used for the evaluation of the decay slope. Linear regression produces a numerically optimal linear estimate, which does not suffer from subjective judgement or loss of accuracy. It is simple to implement on a computer, especially if the the discrete impulse response data already exists in the system.

Converting levels to a decibel scale, the decay can be described by a linear equation $y = ax + b$, where the decay curve of slope a is at level y at time x. Offset b is usually equal to zero, as the curve is commonly normalized to begin at a level of 0 dB, thus passing through the origin. The best-fit slope and bias values can be evaluated in a least-squares fashion.

† An act shrewdly known as "*eyeball integration*" in the practicioners' jargon.

The quality of the line fit estimate produced by linear regression can be described by the correlation coefficient r:

$$r_{xy} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{\left(\overline{x^2} - (\bar{x})^2\right)\left(\overline{y^2} - (\bar{y})^2\right)}} \tag{2.43}$$

The correlation coefficient has the range [–1,1], with high correlation producing values close to unity.

The reverberation decay can be calculated from the impulse response as

$$L(t) = 10\lg\left(\frac{\int_t^\infty h^2(\tau)d\tau}{\int_0^\infty h^2(\tau)d\tau}\right)[\text{dB}] \tag{2.44}$$

This process is commonly known as *Schroeder integration*. Its expected value is

$$L_{\text{expected}}(t) = 10\lg\left(e^{-kt}\right). \tag{2.45}$$

Schroeder [Schroeder 1965] states a relationship between the ensemble average of individual decay curves and the corresponding impulse response:

$$\left\langle y^2(t)\right\rangle_e = \int_t^\infty |h(\tau)|^2 d\tau = \int_0^\infty |h(\tau)|^2 d\tau - \int_0^t |h(\tau)|^2 d\tau \ , \tag{2.46}$$

where $y(t)$ are the decay curves and $h(\tau)$ is the impulse response. This implies that the ensemble average of all individual decay curves can be acquired in a single measurement of the impulse response.

Practical considerations on time windowing

As real impulse responses always contain some background noise, the decay must be approximated using some fixed upper time limit for the integration interval:

$$L(t) = 10\lg\left(\frac{\int_t^{T_i} h^2(\tau)d\tau}{\int_0^{T_i} h^2(\tau)d\tau}\right)[\text{dB}]. \tag{2.47}$$

The upper time limit T_i should be chosen in accordance with the available signal to noise ratio of the impulse response.

Due to the presence of background noise in practical measurements, the slope of the decay curve cannot be evaluated fully in accordance with the definition. The choice of T_i naturally also limits the available lower level for the evaluation of the decay slope. According to ISO 3382, T_i should be selected so that the impulse response is truncated 5 dB above the noise floor.

In order to separate reverberation from the effects of direct sound and strong early reflections, the first 5 dB of the decay curve are excluded for the evaluation of reverberation time. For calculating T_{30} with the standard evaluation interval of –5 dB to –35 dB, this would require a dynamic range of at least (5 dB + 30 dB + 5 dB) = 40 dB, which may be difficult to attain in practice especially at low frequencies.

Many sources such as Gade [1989] and the MLSSA system [Rife 1996] also use a lower limit of –25 dB for the evaluation interval. This yields the index T_{20}, an estimate of T_{30}, that is widely used in practice due to its more relaxed requirements for dynamic range.

2.3.3 Early decay time

Early decay time (EDT) is defined as the time interval required for the sound energy level to decay 10 dB after excitation has stopped [Jordan 1970]. To enable direct comparison with reverberation time, the result is scaled by a factor of 6. For an ideal exponential decay in a diffuse field, the expected value of the EDT equals T_{60}.

Determination of T_{10} is performed in a similar manner with T_{60}, substituting 0 dB and –10 dB for the evaluation interval limits.

EDT takes into account the subjective importance of the early part of the reverberation process [Barron 1995]. A high EDT value indicates much reverberance and low clarity, and vice versa [Gade 1989]. Due to local differences in early reflections, measured EDT values have greater variance than corresponding T_{60} values. Also, the short integration period makes the calculation of EDT sensitive to the accurate determination of the direct sound.

2.3.4 Energy parameters

A number of room acoustical criteria are derived from the energy-time distributions of squared and time-integrated impulse responses. Most of these quantities are of the form

$$C_{t_e} = 10\lg\frac{\int_0^{t_e}|h(t)|^2 dt}{\int_{t_e}^{\infty}|h(t)|^2 dt}\,[\mathrm{dB}]. \tag{2.48}$$

The time limit t_e divides the impulse response into early and late sound energy. Early sound consists of the direct sound and early reflections, whereas the late sound contains the reverberation.

In practice, the upper limit of the integral in the denominator is limited by the relevant impulse response length. Depending on the selection of the time limit t_e (usually defined in milliseconds), the energy ratios emphasize different aspects of the impulse response.

Common time-energy parameters are described in the following.

Clarity

Clarity C is defined as the logarithmic ratio of an impulse response's energy before time t_e, and the energy after t_e, where time t_e equals 50 or 80 ms. The 50 ms value is considered to be best suited in correspondence to the clarity of speech, whereas 80 ms is suited for music.

$$C_{50} = 10\lg\frac{\int_0^{50\,\mathrm{ms}} p^2(t)\,dt}{\int_{50\,\mathrm{ms}}^{\infty} p^2(t)\,dt}\ [\mathrm{dB}] \tag{2.49}$$

$$C_{80} = 10\lg\frac{\int_0^{80} p^2(t)\,dt}{\int_{80}^{\infty} p^2(t)\,dt}\ [\mathrm{dB}] \tag{2.50}$$

High values for clarity indicate large amounts of early energy, which corresponds to a subjective sensation of clarity. On the contrary, low clarity values indicate an unclear, excessively reverberant sound.

Under ideal conditions, the expected value for clarity is given by [Jordan 1980]

$$C_{\mathrm{expected}} = 10\lg\left(e^{kt_e} - 1\right)\ [\mathrm{dB}], \tag{2.51}$$

using the constant k defined in equation 2.36.

Definition

Definition D is defined as the logarithmic ratio of an impulse response's energy before time t_e, and the total response energy, where time t_e equals 50 or 80 ms. The 50 ms value is considered to be best suited in correspondence to the definition of speech, whereas 80 ms is suited for music.

$$\begin{aligned} D_{50} &= 10\lg\frac{\int_0^{50\,\mathrm{ms}} |h(t)|^2\,dt}{\int_0^{\infty} |h(t)|^2\,dt}\ [\mathrm{dB}] \\ D_{80} &= 10\lg\frac{\int_0^{80\,\mathrm{ms}} |h(t)|^2\,dt}{\int_0^{\infty} |h(t)|^2\,dt}\ [\mathrm{dB}] \end{aligned} \tag{2.52}$$

Definition is a measure of the distinctness and clarity of speech and music. It is also commonly known by its German name *Deutlichkeit* [Thiele 1953].

The expected value for D under ideal conditions is

$$D_{\mathrm{expected}} = 10\lg\left(e^{-kt_e}\right)\ [\mathrm{dB}] \tag{2.53.}$$

Center time

Center time corresponds to the center of gravity of the impulse response energy:

$$T_S = \frac{\int_0^{\infty} t\,h^2(t)\,dt}{\int_0^{\infty} h^2(t)\,dt}\ [\mathrm{s}] \tag{2.54}$$

The value of T_S is usually stated in milliseconds. A low value indicates that most of the energy arrives early, while a high value indicates that the energy arrives late after the direct sound. A low T_S suggests a sensation of clarity, whereas a high T_S suggests a reverberant sound. As T_S is very highly correlated with the EDT [Gade 1989], [Gade, Rindel 1984], it seldom contains any additional information when compared to the EDT.

In practice, the upper integration limit is set by the length of the measured response. As the late part of an exponential decay contains a very small fraction of the total energy, it is common to evaluate only the first full second of the response for TS. [Gade 1989]

The subscript S stands for the German name *Schwerpunktzeit* [Kürer 1969].

For an ideal system, the expected value of T_S is

$$T_{S,\text{expected}} = \frac{T_{60}}{13.6}. \tag{2.55}$$

2.3.5 Strength

Strength is defined as the logarithmic ratio between the total impulse response energy and the energy of the direct sound measured 10 meters from the source. For measurement convenience reasons, the direct sound is often determined from the distance of 1 meter and scaled appropriately:

$$G = 10\lg\left(\frac{\int_0^\infty h^2(t)dt}{\int_0^\infty h_{ref,\,1m}^2(t)dt}\right) + 20\,[\text{dB}], \tag{2.56}$$

Another method is to use a stationary calibrated sound source, the level of which has been measured at a distance of 10 meters in a free field.

Strength (sometimes also called level and denoted with L) describes the influence of the room on the perceived sound level. Gade [Gade 1989] has proposed, that the frequency variation of level could be used instead of T_{60} or EDT to measure the influence of the room on timbre or tone color.

The change of G over a distance in a room gives some indication of how diffuse the room's sound field is. The expected value according to Sabine diffuse field theory is given by

$$G_{\text{expected}} = 10\lg\left(\frac{T_{60}}{V}\right) + 45\,[\text{dB}], \tag{2.57}$$

where T_{60} is given in seconds, and the volume V in cubic meters.

2.3.6 Stage parameters

The acoustical quantities and properties at concert hall audience and stage areas vary from one another. An ensemble of musicians playing on stage has a stringent requirement for clarity and support, in order to properly hear one another playing. The distances between sources and receivers are also of a magnitude smaller on stage than at the audience areas.

This sets requirements for a subset of room acoustic parameters, specifically aimed at stage acoustics. Several common stage acoustic parameters are described in the following.

Stage clarity

A *stage clarity* parameter measurered at a distance of 1 meter from the source onstage is defined separately:

$$C_S = -10\lg\frac{\int_{80\text{ ms}}^{\infty} p^2(t)dt}{\int_{0\text{ms}}^{80\text{ms}} p^2(t)dt}\ [\text{dB}] \approx -10\lg\frac{\int_{80\text{ms}}^{\infty} p^2(t)dt}{\int_{\text{direct}} p^2(t)dt}\ [\text{dB}] \tag{2.58}$$

The short distance between source and receiver makes it possible to denote the first 80 ms as consisting of direct sound. Thus with an inverted sign, the clarity measure C_S can be used as a measure of the reverberation level on stage, and it has been found to correlate with musicians' impression of reverberance. It is also known by the name ST_{late} [Gade 1992].

Support factors

The *support factors* ST_{early} and ST_{total} are stage parameters, defined as logarithmic ratios between the early reflection energy sent back to the platform and the energy of the direct sound [Gade 1992]. The measurements are performed at a distance of 1 m from an omnidirectional source onstage.

$$ST_{\text{early}} = 10\lg\frac{\int_{20\text{ ms}}^{100\text{ ms}} p^2(t)dt}{\int_{0\text{ ms}}^{10\text{ ms}} p^2(t)dt}\ [\text{dB}], \tag{2.59}$$

$$ST_{\text{total}} = 10\lg\frac{\int_{20\text{ ms}}^{1000\text{ ms}} p^2(t)dt}{\int_{0\text{ ms}}^{10\text{ ms}} p^2(t)dt}\ [\text{dB}] \tag{2.60}$$

The difference between ST_{early} and ST_{total} is the upper time limit of the nominator. ST_{early}, originally defined by the name ST_1 [Gade 1982], is suggested for measurement of musicians' possibility to hear each other on the orchestra platform. ST_{total}, originally known as ST_2, correlates well with musicians' general judgement on acoustic quality and how well the room supports and assists an individual musician playing. [Gade 1989]

2.3.7 Lateral energy fraction

Lateral energy fraction (LEF, LF, LFC) is the ratio between the energy of early reflections arriving from lateral directions and the energy of direct sound and early reflections arriving omnidirectionally [Barron, Marshall 1981]. The formal definition of LFC is

$$L_{\text{fc}} = \frac{\int_{0\text{ms}}^{80\text{ms}} h^2(t)\cos(\theta)dt}{\int_{0\text{ms}}^{80\text{ms}} h^2(t)dt}\ . \tag{2.61}$$

The lateral energy in the nominator is weighted directionally by a cosine function, where the null axis for the lateral angle θ is pointed towards the sound source.

In practice, an alternate form is used for LF:

$$L_f = \frac{\int_{5\text{ms}}^{80\text{ms}} h_{\text{fig8}}^2(t)\,dt}{\int_{0\text{ms}}^{80\text{ms}} h_{\text{omni}}^2(t)\,dt} \tag{2.62}$$

In order to ensure the elimination of direct sound from the lateral energy, the lower integration limit of the nominator is set to 5 ms after the arrival of the direct sound.

The lateral sound is captured using a figure-of-eight microphone, with the axis placed in the horizontal plane, perpendicular to the direction of the direct sound propagation. This setup can be thought to roughly imitate a listener's ears. The omnidirectional microphone is located as close as possible to the other microphone in order to assure an accurate spatial estimate of the sound field energy.

LEF is a measure of the perceived spaciousness, or the degree of sensation of sound envelopement. High values indicate high spaciousness and vice versa. The parameter has been introduced by Barron [Barron, Marshall 1981].

Another parameter related to LEF is the *lateral efficiency* LE [Jordan 1980], which is defined as

$$LE = \frac{\int_{25\text{ms}}^{80\text{ms}} h_{\text{fig8}}^2(t)\,dt}{\int_{80\text{ms}}^{\infty} h_{\text{omni}}^2(t)\,dt}. \tag{2.63}$$

The expected value of LE and LEF are

$$LE_{\text{expected}} = \frac{2}{\pi}\frac{e^{-t_{\text{lower}}k} - e^{-t_{\text{upper}}k}}{1 - e^{-t_{\text{upper}}k}}, \tag{2.64}$$

where t_{lower} and t_{upper} are the integration limits of the nominators. The expected values are constant under ideal diffuse field conditions.

Practical considerations

According to the formal definition of LEF, the directional cosine weighting should be applied directly to the energy of the impulse response. However, figure-of-eight microphones commonly exhibit an approximately cosine weighted directional response to sound pressure. This causes a contradiction with formal definition, as squaring the cosine weighted sound pressure results in a squared cosine weighting of the energy, effectively narrowing the directional pattern.

Although intricate methods of measuring lateral energy with a true cosine directional pattern have been described in literature [Kleiner 1989], the squared cosine form has gained common acceptance. As a majority of published LEF values use the squared cosine form, it would be difficult to compare results acquired in accordance to the strict definition.

2.3.8 Interaural cross-correlation

Binaural measurements are a method for measuring signal pairs corresponding to the ears† of a listener [Ando 1985]. Impulse response pairs at the ears are acquired using either a binaural head, or an actual listener rigged with in-the-ear microphones.

The *interaural cross-correlation function* IACF is a binaural measure of the difference in the sounds at the two ears, produced by a sound source on the stage. It is defined as [Hidaka et al. 1995]

$$\Phi_{l,r}(\tau)=\frac{\int_{t_1}^{t_2} p_l(t)\,p_r r(t+\tau)\,dt}{\sqrt{\int_{t_1}^{t_2} p_l^2\,dt \int_{t_1}^{t_2} p_r^2\,dt}}, \tag{2.65}$$

where the subscripts l and r designate the sound pressures measured at the left and right ears.

The maximum possible value of the IACF is unity, resulting from identical signals arriving synchronously in both ears.

The variable τ denotes the time difference between the ears. The values for IACF are usually calculated in the range of (-1 ms $< \tau <$ 1 ms) , which corresponds to the maximum natural sound delay between the ears.

In order to obtain a single number measuring the maximum similarity of sound arriving at the ears within the time integration limits and the range of τ, the parameter *interaural cross-correlation coefficient* IACC is defined as

$$\mathrm{IACC}_t = \max\left(\left|\Phi_{l,r}(\tau)\right|\right), \quad -1\ \mathrm{ms} < \tau < 1\ \mathrm{ms} \tag{2.66}$$

There are several standard IACC parameters with different integration periods [Beranek 1996]:

$$\begin{aligned} &\mathrm{IACC}_{\mathrm{A}} && t_1 = 0\,\mathrm{ms} && t_2 = 1000\,\mathrm{ms} \\ &\mathrm{IACC}_{\mathrm{E(arly)}} && t_1 = 0\,\mathrm{ms} && t_2 = 80\,\mathrm{ms} \\ &\mathrm{IACC}_{\mathrm{L(ate)}} && t_1 = 80\,\mathrm{ms} && t_2 = 1000\,\mathrm{ms} \end{aligned} \tag{2.67}$$

The $\mathrm{IACC}_{\mathrm{Early}}$ parameter is a measure of the *apparent source width* ASW, and the $\mathrm{IACC}_{\mathrm{Late}}$ parameter is a measure of the *listener envelopment* LEV.

The time scale is fixed to the arrival of the direct sound at t = 0 ms.

† Microphones are commonly placed at the ear canal entrances in order to acquire the directional effects of the pinnae as well as the head and torso.

2.4 Special applications involving acoustic impulse responses

In addition to standard room acoustical measurements, there are a number of special applications which can be based on acoustical impulse response measurements.

2.4.1 STI and RASTI

STI (Speech Transmission Index) and its common variant RASTI (Rapid STI) are acoustical criteria describing the clarity of speech transmission in an acoustical space. Both criteria are based on the *complex modulation transfer function*, which may be calculated from an impulse response. [Schroeder 1981], [Rife 1992]

2.4.2 Multichannel applications

Most multichannel applications are based on either intensity techniques [Fahy 1989], or techniques applying microphone rows or matrices [Lahti 1990]. The two-channel intensity techniques yield accurate knowledge of the sound field at a single point. Multichannel techniques involving a larger number of microphones can give accurate directional or local information of the sound field.

Sound intensity methods

Common multichannel methods are based on approximations of sound intensity, using two or more receiver points. Basic intensity measurements require the sound pressure information from only two points, but the method may be extended to receiver arrays or matrices for examining the sound field in multiple dimensions.

An important example of basic two-channel intensity applications is the impedance tube used for material measurements [Brüel&Kjær 1992].

Three-dimensional measurements

[Farina, Tronchin 1998] describe a method for performing three-dimensional impulse response measurements of the sound field. The method is based on an array of receiver points situated at the center and end points of a three-dimensional cross. The method enables the determination of impulse responses and sound intensity components for all three Cartesian coordinate axes.

A largely similar approach has been taken in the current TAKU/BINA project using a three-dimensional intensity probe with two individual microphone pairs situated on each Cartesian axis. The probe enables the simultaneous measurement of all six degrees of freedom in a sound field, with separately spaced microphone pairs on each axis for high and low frequency bands.

3 Measurement and analysis methods

This chapter deals with a discussion on various methods and approaches for measuring, processing, filtering and analysing impulse responses. The chapter ends with a discussion on the differences in results between different measurement techniques.

3.1 Impulse response measurement methods

There exist various methods that are suited for the measurement of either the frequency response function or the impulse response function of an acoustical system. These methods may be categorized by the type of excitation.

Impulses

The impulse excitation method is simple and elegant, but suffers from several problems. Since perfect Dirac delta impulses are impossible to generate, devices such as start pistols and spark gaps are commonly used [Bradley 1986] to produce short wide-band impulses with adequate energy to overcome signal-to-noise problems. The absolute level, frequency spectrum and directivity of these sources are prone to variations. Alternatively the impulses may be produced electrically, filtered, and reproduced by a loudspeaker. This decreases the variations, but usually also limits the available signal energy. The slow rise times of energy-time envelopes related to large acoustic spaces may result in only a portion of the room modes to be excited by a short stimulus, yielding an incomplete picture of the acoustical system. Impulse excitation is well suited for the analysis of early reflections, but it is not recommended for the accurate measurement of reverberation times [ISO 3382, 1997], and may produce unreliable results in energy-time parameters associated with long decay times.

Modulated pure tones (warble tones)

Warble tones are an advanced variant of static sinusoid excitation. The stimulus consists of a sinusoid which is frequency modulated around a center frequency. Due to the spread spectrum energy, warble tones are less prone to variances caused by individual room modes at point frequencies, resulting in smoother decay curves [Kutruff 1979]. The frequency response acquired by warble tones corresponds to a filtered impulse response. For wide-band results, an extensively large number of measurements is required.

Slow frequency sweeps (TDS)

Time-Delay Spectrometry (TDS) [Heyser 1988] is a special method involving a sinusoid excitation, which is swept over the frequency range under measurement. The response is filtered with a bank of steep band-pass filters synchronized with the stimulus, producing an excellent signal-to-noise ratio. The TDS method is not limited to LTI systems, but can also be used to measure

strongly nonlinear systems. TDS produces a frequency response function, from which the impulse response may be calculated.

Stationary stochastic signals (noise)

Noise excitation is suitable for measuring either a wide-band or a band-limited frequency response. The stochastic nature of the stimulus causes variance in short-time measurements, and requires the averaging of multiple measurements at each source-receiver location. The possibility of band-limiting can give a good signal-to-noise ratio even with a moderate loudspeaker [Kutruff 1979]. Noise excitation measurements result in a system frequency response.

Synchronized periodic signals (MLS, Schroeder)

The impulse and MLS stimuli exhibit a unity autocorrelation, which enables impulse response measurements directly in time domain. The other excitation types are used to measure the frequency response, from which the corresponding impulse response can be calculated by the inverse Fourier transform.

The properties of MLS measurements have been thoroughly discussed in chapter 2. Briefly, the MLS method offers a fast and accurate wide-band impulse response with an excellent noise tolerance for systems that are reasonably linear and time invariant.

Legendre sequences [Schroeder 1979], Barker codes [Kutruff 1979] and Golay codes [Golay 1961] are examples of other signals exhibiting a unity autocorrelation. Impulse responses are calculated using deconvolution for these signals.

Two-channel methods (FFT/IFFT)

Two-channel FFT methods enable the use of an arbitrary excitation signal, which is cross correlated with the system response in order to yield the frequency response [Bendat, Piersol 1980]. In comparison with single channel methods employing a deterministic stimulus, two-channel FFT enables a true measurement of the acoustic system, without any loudspeaker-induced artefacts included in the results.

3.1.1 Alternative methods for room decay measurement

There are several alternative methods for the measurement of reverberation times in rooms. One group of methods is based on an interrupted noise stimulus, while others use impulses either in direct form (pistol shots, spark gaps, balloons) or in a more complex form (MLS method utilizing pseudorandom noise).

Interrupted noise methods

The acoustical system is first excited using filtered or unfiltered impulses or random noise, which is suddenly cut off [Chu 1978]. The room decay is then filtered and measured using either an analog level detector such as a plotter, or time-averaged and sampled with a relatively slow sample rate in the order of 1–10 ms [Rasmussen, Hansen 1990].

The decay curves of time-varying systems may be accurately determined by averaging over a large number of decay curves measured using the interrupted noise method. Band-filtered noise may be employed to increase the effective signal-to-noise ratio.

However, the use of interrupted random noise has several drawbacks when compared to impulse response measurements:

Decay curves measured in room enclosures exhibit random fluctuations resulting from interference of normal room modes of different natural frequencies. Amongst various other parameters, the exact form of the random fluctuation depends on the initial amplitudes and the phase angles of the normal modes at the moment that the excitation signal is turned off. For a random noise stimulus, these parameters are different from trial to trial, resulting in an amount of randomness in the resulting decay curves. With static room geometry and transducer locations, the excitation signal is the main cause for randomness in the results.

Traditionally a number of repeated measurements has been made, from which the results have been calculated as a statistical mean value. As the statistical error term is related to the number of measurements squared, a large number of measurements has been needed for accurate results.

In addition to practical inefficiency, this method is unable to reveal smaller details of the decay curve, such as multiple decay rates and high initial rates of decay, which can be of prime importance in evaluating the decay quality in a room or hall.

For comparison, the reverse-time integration of an impulse response (*i.e.* Schroeder integration, see Chapter 2) results in a decay curve identical to the average of infinitely many decay curves measured by the interrupted noise method is available in a single measurement.

Pure tones (sinusoids)

Pure tones are sometimes used for investigating individual widely spaced low-frequency room modes. Decay curves resulting from sinusoidal excitation show an uneven characteristic due to the strong local coupling of individual room modes. [Beranek 1988]

3.2 Acquisition, processing and analysis methods

The following methods provide improvements to the measurement, processing and filtering of impulse responses by taking into account and minimizing the effects of certain sources of error.

3.2.1 Methods for enhanced measurements

Several signal processing methods are commonly used to enhance practical measurements of impulse responses. These include filtered stimuli, block averaging and response compensation.

Filtered stimuli

Filtered wide-band stimuli can be better suited for practical measurements than common flat-spectrum signals, such as the MLS or white noise. In order to obtain a wide-band impulse response, inverse filtering is applied to the acquired response.

Stimuli with a pink-shaped (–3 dB/octave) spectrum are commonly used for significant gains in the signal-to-noise ratio. This is due to the following two factors: The power handling capacity of high-frequency loudspeaker units generally limits the available signal levels for wide-band stimuli. The pink filter reduces the amount of power required for the high frequencies, enabling for a substantial increase in gain at the low frequencies. Secondly, most practical measurements are inevitably influenced by relatively high low-frequency background noise levels, due to the amount of low-frequency noise sources in the environment. Thus, a desirable improvement in the low-frequency signal-to-noise ratio is achieved.

Block averaging

Block averaging is used to calculate the average response from a number of successive response cycles. With linear operations used for the recovery of the impulse response, such as for the MLS method, block averaging can be employed either before or after the calculation of the impulse response. Either method exhibits certain advantages and drawbacks.

With *pre-averaging*, only a single impulse response calculation is required, as the raw response cycles are averaged together prior to the calculation. Pre-averaging can reduce the storage required for the input data to a single cycle, if successive input cycles are averaged during the measurement. This method enables infinitely long averages to be taken, assuming that the system under test is highly linear and time-invariant.

With *post-averaging*, an impulse response is calculated for each raw response cycle prior to averaging. Post-averaging can be a valuable tool for evaluating and possibly discarding certain impulse responses before averaging. Applications of post-averaging include noisy and unstable conditions, where a long average is required, but many of the measurement cycles are corrupted by

noise. Post-averaging can also be used with care in slowly time-variant conditions, where an optimal time-average estimate may be obtained by studying the rate of change between the successive impulse responses. [Nielsen 1996], [Nielsen 1997]

Response compensation

Any linear systematic errors included in the measurement signal chain can be effectively compensated for by deconvolving the acquired responses with a reference response. The reference response may include one or several of the following factors to be compensated for in the signal chain:

The electrical loopback response of the measurement system can be used to remove the ringing effects cause by anti-alias filtering in the AD converters, as well as most amplitude and phase errors in the system's frequency response.

An acoustical loopback response can be used to compensate for loudspeaker or room response effects, although time windowing may be required.

Another possibility is the use of inverse filtering of responses acquired with a spectrally colored stimulus.

3.2.2 Processing of impulse responses for room acoustical analysis

Real measured acoustical impulse responses can greatly differ from their ideal theoretical counterparts, which portray a noiseless and truly exponential decay process. As the definitions of most room acoustical parameters are based on these ideal models, practical impulse responses are seldom directly applicable for calculation, if accurate results are desired. Thus real responses must first be "cleaned up", in order to avoid calculation problems due to unideal properties.

The main differences between practical measured impulse responses and ideal impulse responses are threefold: 1) A real response exhibits some length of initial delay before the arrival of the direct sound. 2) The decay may be non-exponential, or it may consist of multiple parts with different slopes. 3) The response contains background noise, which limits the decay to a noise floor at some point.

These effects require specific attention in order to calculate values for room acoustical parameters with reasonable accuracy and repeatability. Neglecting this can result in excessive and varying values for the EDT and T_{60}, and a large degree of randomness in the energy-time parameters.

A detailed case study is given in Chapter 5.

Determination of direct sound and the background noise level

Due to system propagation delays, the direct sound does not arrive exactly at the beginning of the measured impulse response, but at some later point depending on the measurement conditions. A good estimate for the initial delay can be calculated from the source-receiver distance, if available. This method is rather unpractical for extensive field measurements of large spaces, such as concert halls. Even if an accurate distance is known, other factors such as the

ambient temperature and speaker crossover delays cause variations to the initial delay actually measured. This calls for methods for the determination of the initial delay directly from the measured impulse response.

The ISO 3382 standard gives the following guideline: "[The starting point for direct sound] should be determined from the broadband impulse response where the signal first rises significantly above the background but is more than 20 dB below the maximum." In addition, the standard suggests that this result be compensated for use with filtered responses according to the individual filter delays. [ISO 3382, 1997]

A traditional method is to rely on the system operator to visually determine a starting point from a graph of the measured impulse response. This enables the operator to use his expertise and judgment, but causes some variance in the results between different operators and occasions of evaluation. With a large number of measurements, this method also calls for considerable time and effort.

Alternatively, an algorithm may be devised to determine the initial delay automatically. This ensures identical results for a certain response, but may fail to correctly evaluate difficult cases. In order to comply with the ISO guideline, a number of parameters have to be evaluated, *i.e.* the maximum level, the background noise level, and a level for the starting point. Determining the maximum level of the squared impulse response is a trivial task, but the estimation of the background noise level is somewhat more complicated. In order to determine the noise level, it is necessary to find a section of the response dominated by noise, and calculate the RMS value. This requires certain knowledge on the temporal structure of the response, in order not to account for a late part of the decay as noise. Finally, an appropriate level for the starting point should be determined from the two previous values.

Determination of the decay-noise crosspoint

The limited signal-to-noise ratio inherent in practically all acoustical measurements, and especially measurements performed under field conditions, calls for attention concerning the evaluation of the Schroeder time integral's upper time limit. Theoretically this limit is set to infinity, but in practical measurements it is naturally limited to the length of the measured impulse response data. In practice, measured impulse responses must be long enough to accommodate for the whole room decay down to the background noise level, in order to avoid time aliasing (MLS and other cyclic impulse methods). Thus the measured impulse response contains not only the decay curve under analysis, but also a steady level of background noise, which dominates for some length of time at the end of the response. Integrating this steady energy level along with the exponential decay curve causes an error both in the resulting decay factor T_{60} and in the time-window energies (*i.e.* energy parameters).

In order to avoid the effects of the noise artifacts, analysis must be performed on the impulse response data in order to find the level of the background noise, and a point where the room decay meets the noise level. This way it is possible to effectively truncate the impulse response at the start of the noise level, minimizing the noise energy mixed with the actual decay curve.

In addition to determining the starting point, it is thus essential to find an end point where the decay curve meets background noise, and truncate the noise from the end of the response. According to the ISO 3382 standard, the noise floor must be 10 dB below the lowest decay level used for calculation of decay slope.

Similar methods can be used for determining the end point as for the starting point. If the noise floor level can be estimated with good precision, it is simple to search for the first point on the decay curve at a level of 5 dB above the noise floor, proceeding backwards towards the start.

The evaluation of the end point has a notable effect on the resulting decay times. If the end point is selected too far into the decay, the truncated integrated impulse response will fall too early, resulting in shorter decay times, and vice versa.

There is no standardized exact method for determining the integration limits and noise compensation for integrated impulse responses. [Lundeby *et al.* 1995] propose guidelines for an iterative algorithm, where the crosspoint of the background noise level and the late decay curve are iteratively estimated from several time-averages of the squared response. This is described in more depth in Chapter 5.

[Faiget *et al.* 1996] propose a simple but systematic method for post-processing noisy impulse responses. Assuming that the acquired response is longer than the real duration of the room decay, the latter part of the response is used for the evaluation of a background noise level by means of a regression line. In order to avoid strong and discrete early reflections, the decay starting point is determined corresponding to a multiple of the mean free path between the source and the receiver. The decay ending point is chosen at 5 dB above the noise floor, and the end of the useful response is determined at the crosspoint of the decay and background noise regression lines.

Compensation for truncated decay energy

The truncation of the noise tail from an impulse response also cuts off the remaining portion of the decay curve that is left under the noise floor. When the resulting decay curve is Schroeder integrated, it falls of steeply at the point of truncation. This systematic error makes the latter part of the decay slope steeper, and may result in underestimated reverberation times.

[Lundeby *et al.* 1995] also describe a method for estimating the amount of lost decay energy. The actual decay curve is artificially extended beyond the point of truncation, by extrapolating the regression line on the late decay curve to infinity. The total compensation energy is formed as an ideal exponential decay process, the parameters of which are calculated from the late decay slope.

Decreasing the effects of background noise

A large signal-to-noise ratio is required for the accurate determination of decay slopes. [Hirata 1982] has proposed a simple method for improving the signal-to-noise ratio by replacing the squared single impulse response $h^2(t)$ with the product of two impulse responses measured separately at the same position:

$$\begin{aligned}\int_t^\infty h^2(t)dt &\equiv \int_t^\infty [h_1(t)+n_1(t)][h_2(t)+n_2(t)]dt \\ &= \int_t^\infty [h_1(t)h_2(t)+h_1(t)n_2(t)+h_2(t)n_1(t)+n_1(t)n_2(t)]dt \\ &= \int_t^\infty \left\{ \underbrace{h_1(t)h_2(t)}_{\text{correlated, } \approx h^2} + \underbrace{n_1(t)n_2(t)\left[1+\frac{h_1(t)}{n_1(t)}+\frac{h_2(t)}{n_2(t)}\right]}_{\text{uncorrelated, } \approx \pm n^2} \right\} dt \\ &\approx \int_t^\infty h^2(t)dt + K(t)\end{aligned} \tag{3.1}$$

The measured impulse responses consist of the decay terms $h_1(t)$, $h_2(t)$ and the noise terms $n_1(t)$, $n_2(t)$. The highly correlated decay terms yield positive values corresponding to squared response, whereas the mutually uncorrelated noise terms are seen as a random fluctuation K(t) superposed with the first term.

Unlike the Schroeder method described in Chapter 2, this integrated impulse response does not continuously decrease due to the random variations of noise. This makes it easy to visually distinguish noise from the decay curve.

Under adverse noise conditions, a direct determination of the T_{30} decay curve from the squared and time-averaged impulse response has been noted to be more robust than the integrated impulse method. [Satoh *et al.* 1998]

[Xiang 1995] describes a nonlinear approach for the determination of the regression line on the decay curve, which can yield precise reverberation times without need for careful determination of the integration limits or the background noise level.

3.2.3 Methods for filtering responses

Room acoustic parameters are often measured in standard octave or third octave frequency bands [IEC-1260, 1995]. Using a wideband stimulus for the measurements, it is practical to apply the filters during post-processing of the wideband impulse response.

Digital realization

Digital filtering directly corresponding to the standard analog counterparts may be realized using Butterworth IIR filter coefficients and a direct form II transposed filter [Oppenheim, Schafer 1989].

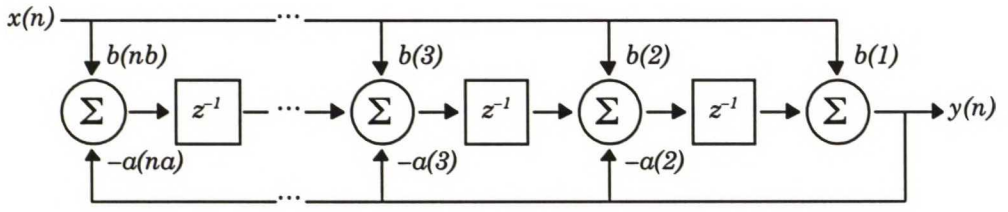


Figure 3.1. Direct form II transposed digital filter structure. [Mathworks 1999]

The operation of the filter for sample n is given by

$$\begin{aligned} y(n) = {} & b(1)x(n) \\ & + b(2)x(n-1) + \ldots + b(nb)x(n-nb+1) \\ & - a(2)y(n-1) - \ldots - a(na)y(n-na+1) \end{aligned} \tag{3.2}$$

where b and a contain the filter specific numerator and denominator coefficients.

Filter delay

Standard octave and third octave Butterworth filters exhibit a causal response, which causes delay and and limits the decay of the filtered signal especially at low frequency bands. These artifacts can cause certain problems in calculating room acoustic parameters from filtered responses.

Filter delays must be taken into consideration upon determining the time of arrival for direct sound in a filtered impulse response. The delays are filter specific, and vary with frequency.

Filter decay not only blurs fine time domain details in impulse responses, but also sets the lower limits for reverberation time measurements involving extremely short decay times, in the order of 100–200 ms. The decay times are also filter specific, varying with frequency. See Chapter 4 for measurements of the filter decay times realized in the system.

Jacobsen and Rindel [Jacobsen, Rindel 1987] set requirements for filter properties for correctly analyzing a given decay rate. For ordinary analog filters (and their similar digital counterparts), acceptable decay curves are obtained only if

$$B T_{60} > 16, \tag{3.3}$$

where B is the filter bandwidth and T_{60} is the reverberation time of the system under test.

Time-reverse filtering

Figure 3.2a depicts the impulse response of a standard 100 Hz third-octave band filter. It is clearly observed that the filter's impulse response is not symmetric in time. The rising part of the filter response is distinctly shorter in time than the "tail" part.

A signal may be reversed in the time domain without effecting its amplitude spectrum. This makes it possible to revert a filter response backwards in time before convolving it with another signal that is to be filtered. Figure 3.2b shows the integrated filter decays in a Schroeder plot. The time-reversed filter decays 60 dB in under 60 ms, while the ordinary filter takes over 200 ms to decay.

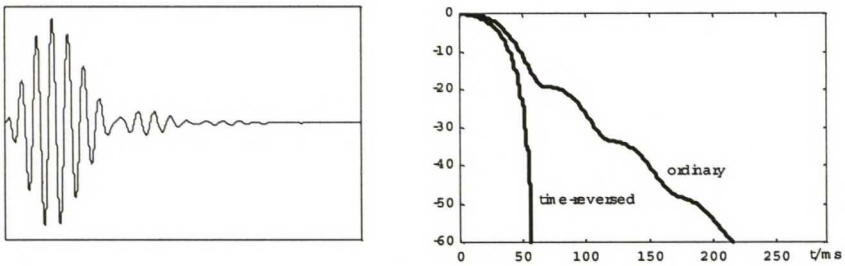


Figure 3.2a. Impulse response of a 100 Hz third-octave band filter.
Figure 3.2b. Schroeder integrated direct and time-reversed filter responses.

This fact enables the effects of filter decay to be effectively reduced by performing the filtering in the time-reversed direction.

The practical realization of time-reversed filtering is simple, when data analysis is performed on a computer. If real-time operation is not required, as with filtering acquired impulse responses, the filter input data is prone to exist in complete form and be of finite length. A practical alternative to designing digital implementations of time-reversed filters is simply to time-reverse the data before and after filtering, and make use of the ordinary filters. From a convolution point of view, it makes no difference whether the filter response or the data is reversed in relation to one another.

For time-reversed filters, Jacobsen and Rindel [Jacobsen, Rindel 1987] give a more relaxed alternative to equation 3.3 as

$$B T_{60} > 4 , \tag{3.4}$$

assuming that the upper 5 dB are excluded from the evaluation of decay curves. This is in good accordance with the T_{20} and T_{30} definitions for reverberation time, as specified in the ISO 3382 standard [ISO 3382, 1997].

According to equations 3.3 and 3.4, the filters depicted in figure 3.2b are thus suitable for filtering reverberation times in the order of 100 ms for the time-reversed filter, and 400 ms for the ordinary filter. Although still limited, the time-reversed filter can be considered adequate for use with all practical room acoustic measurements, whereas the ordinary filter could cause potential problems in small highly absorbent spaces. For even lower frequencies, the filter decay times naturally increase from these values. At this point, however, the primary limit for measurement accuracy is set by the definition of reverberation, which becomes vague with very low frequencies in small spaces.

Time spreading of energy

The blurring effect of low-frequency filters may cause peaks in the signal energy to spread in time. This is undesirable especially with the calculation of

acoustical energy-time parameters, such as C_{80} or C_{50}. Here blurring can cause the possibility of misinterpretation, as the initial energy of the direct sound and the first reflections is spread over the upper time limit of integration and is mistaken as belonging to the late sound.

The ISO 3382 standard recommends to compensate the spreading of energy by extending the upper limit of the early time interval in energy-time parameters by one half of the filter delay time. The energy spread by the filter is thus effectively divided into equal halves for the early and late time intervals.

3.3 Comparison between measurement techniques

There exists a wide variety of measurement techniques and equipment for the measurement and calculation of room acoustic criteria.

Several studies have been carried out for evaluating the variance and differences in room acoustic measurement results between a number of measurement systems and methods [Vorländer, Bietz 1994], [Lundeby *et al.* 1995], [Halmrast *et al.* 1998], [Fausti *et al.* 1998].

MLS-based systems

Impulse responses acquired with different MLS-based measurement systems have been found to be mutually well correlated [Lundeby *et al.* 1995], [Fausti *et al.* 1998]. There are small deviations to be found between systems, such as background noise floor levels and loopback impulse response shapes. These deviations result from the different types of A/D and D/A converters and the analog electronics involved in the system hardware.

Differences between systems are most obvious when an electrical loopback measurement is performed by connecting the system electrical signal output directly to the input. Lundeby *et. al.* [Lundeby *et al.* 1995] performed comparative measurements by using a high-quality band-pass filter unit as the test object.

Alternative methods, such as pistol shots and noise stimuli exhibit a larger variance than MLS-based systems. The amount of variance differs also depending on the type of parameter and frequency band in concern. Reverberation times are generally well correlated, but energy parameters tend to exhibit large deviations especially at low frequencies. Methods employing a random noise stimulus show a considerably higher uncertainty than the impulse methods. In order to achieve comparable uncertainties, at least 3–5 ensemble or T_{60} averages are required [Vorländer, Bietz 1994].

Effects of stimulus type

In the calculation of reverberation time, the major differences are not between stationary and impulsive techniques, but between stationary and impulsive sources [Fausti *et al.* 1998]. Results obtained using the MLS method are closer to the results obtained with stationary interrupted noise than to those obtained by purely impulsive stimuli, such as pistol shots or balloon pops.

Impulse responses measured with pure impulsive stimuli are comparable to results measured with the aforementioned techniques. At low frequencies, the signal-to-noise ratio is commonly lower than with the MLS technique. [Fausti *et al.* 1998]

Differences in room acoustic analysis

The variety of methods employed in the post-processing of impulse responses introduces systematic differences in the resulting values for room acoustic parameters. These differences are caused by the time-windowing, filtering, reverse-time integration and noise compensation procedures, so that two different analysis systems may produce variable results, even when starting from a common impulse response.

The magnitude of the differences between systems has been investigated in literature [Lundeby *et al.* 1995], with deviations of 5–10 % for reverberation times, center time and definition and 0.5 dB for clarity and strength at mid-frequencies. For low frequency bands, major uncertainties can occur between systems. Large variances were found for the lateral fraction (LF) parameter, probably due to the difficulties in determining the sensitivities of figure-of-eight microphone capsules.

For the evaluation of reverberation time, the repeatability of the most accurate impulse and MLS methods has been found to be of the order of 1 %. This is more than adequate for subjective purposes, as only changes of approximately 4 % in reverberation time are found subjectively noticeable [Cremer, Müller 1978, referred to in Vorländer, Bietz 1994]. For measurements in technical and building acoustics, this uncertainty is also negligible when compared to the other factors, such as the spatial variance of a sound field, that cause inevitably larger measurement uncertainties.

At low frequency bands, the MLS method has been found to produce systematically higher reverberation times than impulse and noise methods. It is argued that in this case methods employing band-filtered noise bursts yield the most accurate results. This can be due to a lack of linearity, time invariance or ergodicity in the acoustical and measurement systems. The broad-band MLS stimulus excites a number of narrow and barely overlapping room modes both inside and outside the one-third octave frequency band under test. Vorländer claims that the energy transfer between modes is apparently not invariant with time, causing the minor deviations found between the results obtained by the different methods.

The evaluation of decays from true impulse stimuli is especially prone to error at the low frequency bands, as the decay curves tend to fluctuate much. This is due to the extremely short stimulus, which does not fully excite all room modes.

For the analysis of very short reverberation times, both the interrupted noise method and the integrated impulse response method have been found acceptable. [Rasmussen *et al.* 1991]

4 Implementation of the system

This chapter begins with a briefing on the instrumentation required for an acoustical impulse response measurement system. The rest of this chapter is devoted to the design, coding and implementation of the IRMA measurement system hardware, the Matlab software and the sound card interface program.

4.1 Instrumentation for impulse response measurement

Figure 4.1 depicts a typical computer-based system for measuring the response of an acoustic system. The measurement system is divided into three different signal domains: the acoustic system, the analog electrical system and the digital system. The transfer of signals between the acoustic system and the computer requires analog-to-digital and digital-to-analog converter units, as well as common electro-acoustic transducers, such as loudspeakers and microphones.

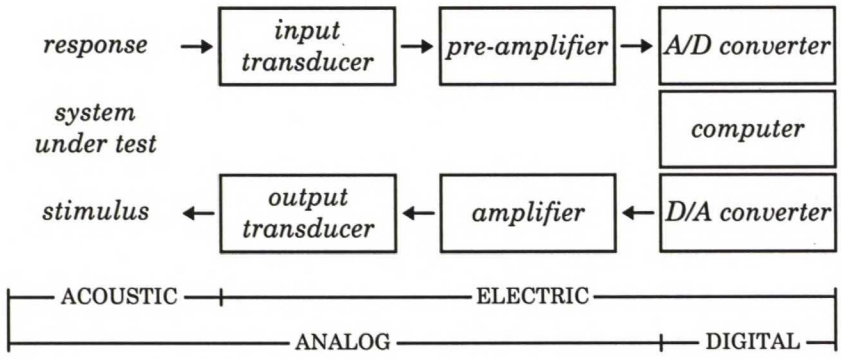


Figure 4.1. The signal path and signal domains in a computer-based system for impulse response measurement.

Measured impulse responses that are evaluated on the measurement computer inherently consists of not only the response of the system under test, but also of all the other system units, *i.e.* the converters, amplifiers and the transducers taking part in the signal path. In order to focus on the system under test, the quality and magnitude of the deviations introduced in the signal chain must be carefully considered and evaluated. This scenario involves a number of parameters to be taken into account for each unit.

As the components are connected in series with one another, any deviations and distortion accumulate in the signal. In a serial configuration, the throughput signal quality is ultimately determined by the lowest quality component in the signal chain.

Due to the nature of the MLS measurement method, the requirements of linearity and time invariance are at premium. The requirements for linearity generally include a linear complex frequency response, as well as low noise and distortion parameters in the signal band. The requirement of time invariance is of interest in the mutual timing and synchronization properties of the analog/digital converter units.

The effects of transducers and analog signal conditioning are discussed more thoroughly in Chapter 6.

4.2 The IRMA system hardware

A portable measurement system was put together as a part of the work. The hardware requirements are discussed next, followed by details of the hardware implementation for the IRMA system.

General requirements

The general system requirements comprised of the following: multichannel operation with adequate inputs and outputs for measurement microphones and sound sources, use of Windows compatible sound hardware for programming reasons, portable but reasonably rugged construction, suitability for field use, and an open architecture enabling for future expansion or modification to suit other applications.

Figure 4.2 shows a basic functional layout for the IRMA system hardware.

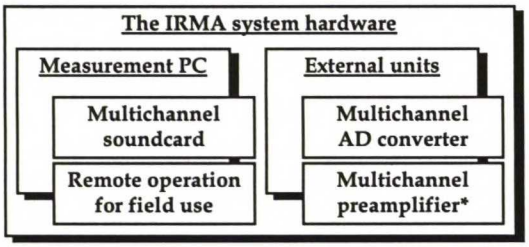


Figure 4.2. A functional layout of the IRMA system hardware.

4.2.1 Audio requirements

The choice of sound hardware was considered of primary importance in the system architecture. The sound card forms a central part of the measuring system, affecting parameters such as measurement accuracy as well as the number of channels and the connections available. In theory, basic measurements could be accomplished with any Windows compatible sound hardware, but this could result in severe limitations in the quality and reliability of the results.

The following audio requirements were set for the measurement system:

- 16 bit or better conversion,
- 44.1 and 48 kHz sample rates,
- linearity, low noise and distortion,
- simultaneous use of multiple input channels and two output channels,
- synchronized timing between channels,
- compatibility with the standard Windows sound device architecture.

The selection of suitable PC sound hardware was first divided by the multichannel requirement together with the need for interchannel synchronization: quality ISA sound cards were only available as two-channel

models due to ISA bus bandwidth limitations, and there was no proper clock sync available to synchronize multiple cards together.

Most industrial data acquisition cards (both ISA and PCI) were also rejected, as they did not offer simultaneous use of multiple input channels using high sample rates. Multichannel data acquisition cards are commonly implemented with a single A/D converter, that is multiplexed between a number of inputs at a high frequency. The time-division scheme involved in the multiplexing tehnique was also feared to cause problematic synchronization issues.

This led to a yet broad selection of PCI sound cards, mostly intended for semi-professional sound studio applications. This seemed to suit well with the aim to realize a measurement application, that was actually independent of the sound hardware by utilizing the Windows Sound API.

Yet only few products offered multiple channel operation using standard audio interfaces combined with standard Windows sound drivers. Product specific hardware/software interfaces were a cause for rejection, as they did not offer methods for writing custom code for using the sound hardware with the measurement application. Several products were also dropped due to proprietary connections to product-specific external analog/digital conversion units, that offered no possibility for interchange or future upgrades to the hardware.

4.2.2 The audio system

Sound card

The Korg 1212 I/O sound card was chosen as most suitable for the application, as it fulfilled most requirements and did not suffer from serious limitations.

The card consists of a PCI interface for adequate data throughput, an 8 channel ADAT/T-DIF optical digital input and output, 2 channel digital coaxial S/PDIF input and output, and a 2 channel analog line input and output. There are also connections for clock sync input and output, so that external digital audio units can be synchronized with the sound card. Details on the card are given in Appendix B.

The sound card is supplied with standard Windows 95/98 drivers, which make the device available as 6 standard two-channel Windows sound devices, all of which can be accessed simultaneously. The interchannel synchronization issue is acknowledged in the device driver software, providing sample-accurate synchronization between all the input and output channels.

External A/D converter

Although there are only 2 analog input and output channels available directly on the sound card, the optical 8 channel ADAT/T-DIF interfaces make it possible to connect external 8 channel A/D and D/A converter units to the card. With the addition of a Korg 880 A/D converter unit to the system, a total of 10 analog input channels are available for simultaneous use. If necessary, the

number of analog output channels may also be increased accordingly with the addition of an external 8 channel D/A converter unit.

The standard ADAT interface makes the system compatible with a large number of different converter units. This gives some choice over the individual properties of the chosen converter unit, such as the connections and the type of A/D converters employed. It is also possible to upgrade from one converter unit to a another model, for example to increase the available bit depth and dynamic range, or to use converters with a more suitable impulse response in the anti-aliasing filters.

The clock sync connection is used to synchronize the external A/D converter unit's sample clock to the sound card's master clock. For special purposes [Mommertz, Muller 1995], it would even be possible to further synchronize the whole system to an external sample clock source using the sound card's sync input. However, this option is not employed in the present application.

Microphone preamplifier

The need of a custom multichannel microphone preamplifier was discussed during the course of this work. No suitable products with fixed gain settings were found to be commercially available. In the future, a custom unit could be designed and built with the following specifications:

- adequate audio quality,
- 8 or more discrete channels,
- phantom power supplies for microphones,
- a choice of balanced/unbalanced connections,
- accurate gain settings with discrete values ranging from 0 to 60 dB.

Optimally the gain settings could be controlled digitally and interfaced directly from the measurement computer, enabling automatic adjustment and storage of the individual channel gains together with the measurement setup.

The phantom supply and the choice of input connectors could offer a range of options, such as 5 V for electret capsules (unbalanced input with BNC connectors), 48 V for standard studio equipment (balanced input with XLR connectors) and even 200 V for laboratory measurement equipment (unbalanced B&K type input with Lemo connectors).

4.2.3 Computer hardware requirements

Several factors affected the choice of components for the system setup. The primary requirement was that the system could house a full length PCI sound card. Other requirements included reasonable computing resources, and such factors as portability and the possibility of field use were taken into account.

A standalone notebook computer would have been the ideal choice regarding the portability issue, but the requirements for good audio quality, multichannel operation and rugged construction could not be fulfilled with a reasonable budget. As the multichannel sound hardware inherently requires a PCI card (suitable USB or Firewire products were not available at the time), a laptop computer would require at least a large docking station for housing the

expansion cards, as well as the external A/D converter and microphone preamplifier units.

Since microphone amplifier and A/D converter units are generally in 19" rack format, it was a natural alternative to house the measurement computer in the same rack. Industrial rack computers are readily available, but unlike laptop computers, the units do not include portable console devices.

Under stationary operating conditions, an ordinary external display and keyboard can naturally be used with the rack computer. For field measurements, a portable solution was still required, as available rack mountable lcd display and keyboard units exhibited excessive weight and price figures.

It was decided to experiment, using a separate laptop computer for operating the measurement system with remote software over a network connection. All data acquisition, processing and storage are still handled by the rack computer, with the laptop only used as a remote mirror of the rack computer's display and input devices. This solution enables using the measurement system remotely from the physical rack unit, with only a single twisted pair network cable of optional length attaching the units. Not only does this make operating the measurement system more comfortable, such as being able to sit in the audience area of a hall instead of kneeling down at a dim stage corner, but it also offers a means for reducing the fan noise level emitted by the rack computer, by freedom of placement and even partial covering of the measurement computer rack.

4.2.4 The computer system

The measurement computer comprises of an Pentium II computer, housed in an industrial rack case and equipped with a multichannel PCI sound card. In addition to the computer unit, the measurement system includes an external 8 channel A/D converter and an optional multichannel microphone amplifier, all of which may be housed in a standard 19" rack flight case. Details on the computer hardware are listed in Appendix B.

A 100 megabit Fast Ethernet network connection was chosen to ensure smooth remote operation. No CD-ROM drive was incorporated in the measurement computer to reduce space and weight, as software installation can be performed over the network connection from another computer with a CD-ROM drive.

Remote operation under field conditions was realized using *Symantec pcAnywhere* remote software, which was installed on both the measurement computer and a separate laptop.

4.3 The IRMA Matlab software

Background

Custom software was written for the measurement system, including a full duplex multichannel sound card handler and functions for creating stimuli, performing measurements, post processing, filtering and analyzing responses, as well as a graphical user interface for operating the main features.

Matlab was chosen for the basis of the signal analysis system, as it enables a wide range of powerful and fast tools for numerical matrix analysis, as well as a programming environment for running simulations and developing applications. The Matlab environment is based on a command line interface from which the workspace contents, scripts and functions are accessible. Matlab version 5 also includes tools for building a graphical user interface. The system can be run in Matlab under Windows 95, 98 or NT. With the exception of the sound card handler, all the application code can be easily ported to other platforms.

Matlab enables the generation and import of arbitrary stimuli, as well as filtering and inverse filtering of stimuli and responses. Along with the multichannel measurement capabilities, these properties make the measurement system suitable for a wide range of applications.

As outlined in Chapter 3, a variety of methods may be used for measuring impulse responses. The IRMA system is focused on the MLS method. It must be emphasized, however, that the system is not tied to the MLS in any way, as impulse response measurements may be performed using any alternative stimuli and FFT deconvolution instead. The MLS method gained focus, because it is widely used in room acoustic measurements due to its high accuracy, speed and applicability to the measurement of room acoustic parameters. There are a number of popular commercial acoustic measurement systems based on the MLS, such as the DRA Labs' MLSSA [Rife 1996], the Aurora system [Aurora 2000] and the WinMLS system [WinMLS 2000].

Application overview

Figure 4.3 shows an overview of the IRMA measurement application software. The application contains Matlab functions for measurement, analysis and display along with a graphical user interface (GUI) and a global measurement setup structure.

Details of the application are discussed in the following sections.

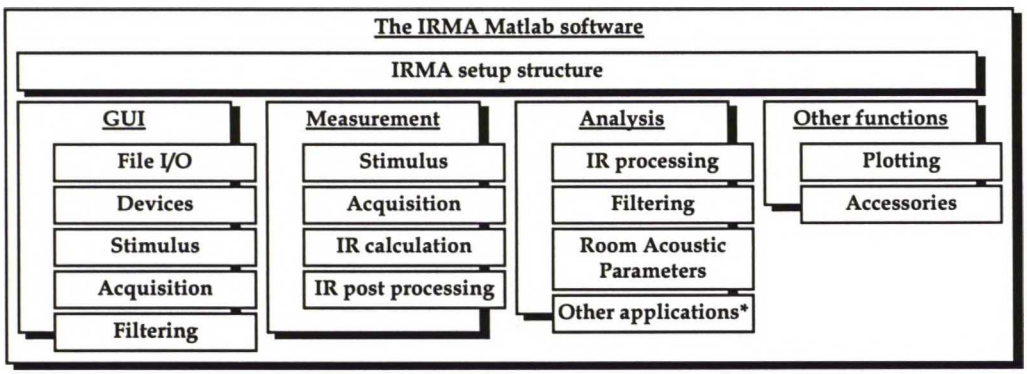


Figure 4.3. A functional layout of the IRMA Matlab software.

4.3.1 Measurement process

The impulse response measurement process can be categorized into four parts: stimulus generation, data acquisition, calculation of the impulse response and post-processing of the response. Figure 4.4, shows a further subdivision of each part, details of which are given in the following sections.

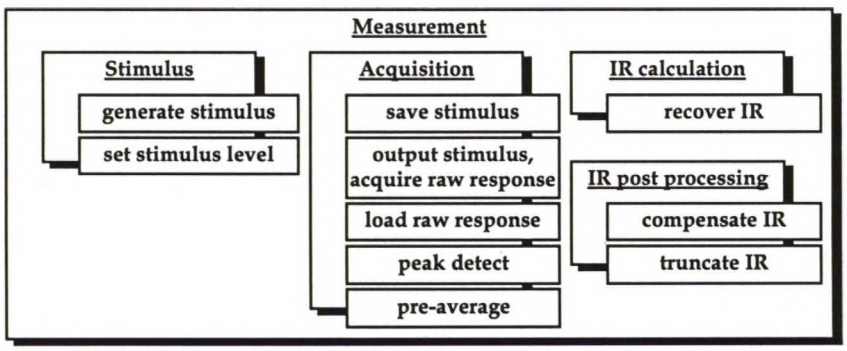


Figure 4.4. An overview of the measurement process.

The implementation of the measurement process is shown in Figure 4.5 on a functional level. This provides a deeper insight into the Matlab functions involved.

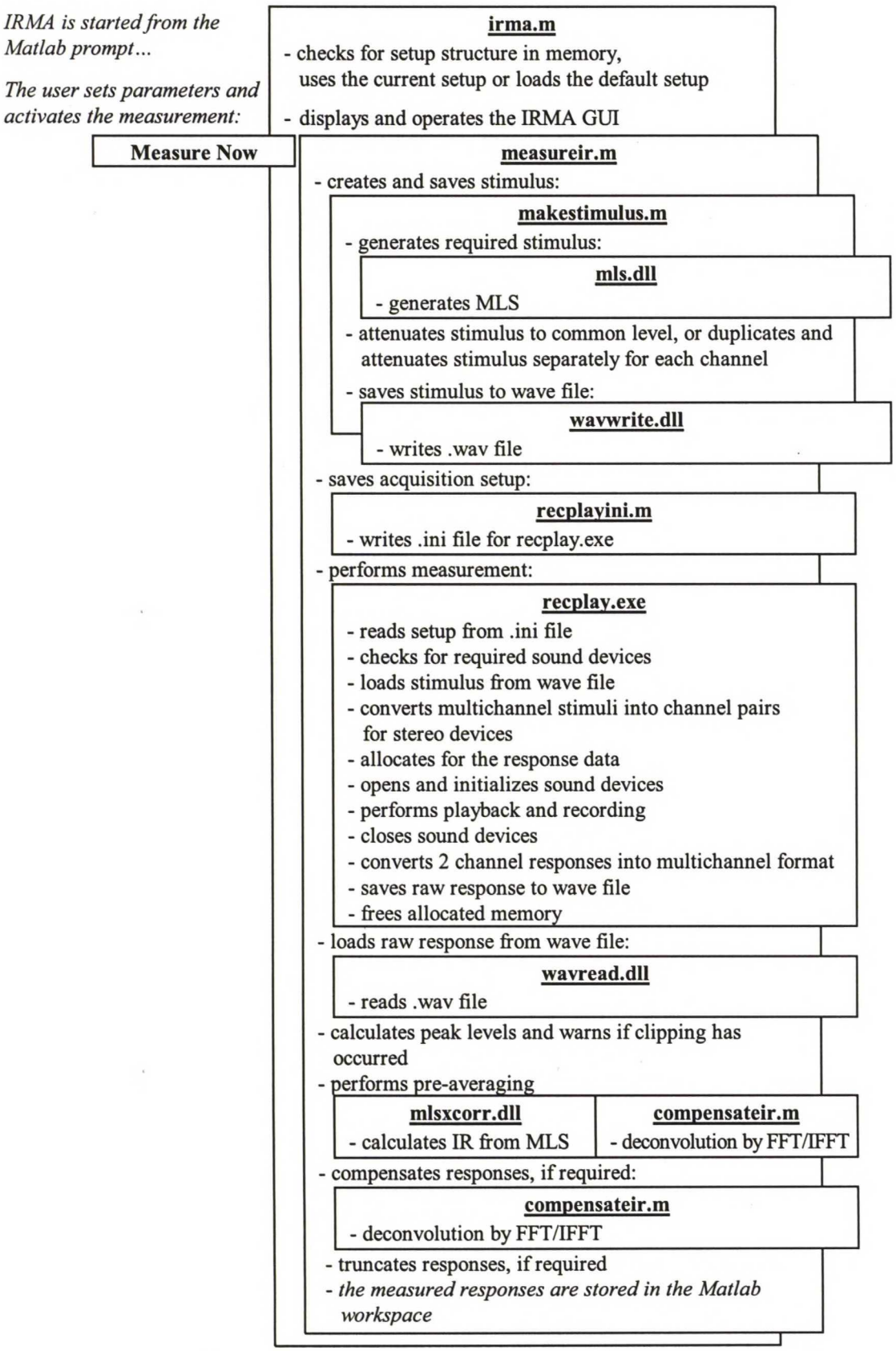


Figure 4.5. The measurement process at function level.

Stimulus generation

First, a stimulus is generated and stored in memory. Current options for stimulus type are maximum length sequences, periodic unit impulses, and arbitrary stimuli supplied by the user.

MLS

MLS stimuli are generated with a binary delay line, as described in Chapter 2. The generation process was realized as a Matlab MEX function written in C, due to speed and memory considerations. Notable increases in speed were gained by realizing the delay line as a single 32 bit unsigned long integer, and applying binary shifts, bit masks and bitwise logic for operating the delay line.

The Matlab environment causes a major inefficiency in memory usage, as the binary sequence output by the MEX function is converted to standard double float format, consuming some 64 bits per binary sample when stored in the Matlab workspace. Matlab version 5 does offer some capabilities for handling alternate number formats, such as 8 bit integers, but they cannot be properly used for calculations. For realistic sequence lengths, the memory requirements can still be handled with normal hardware: a stimulus of length $2^{19}-1$, lasting over 10 seconds with a sample rate of 48 kHz, consumes some 4 megabytes of memory in the Matlab environment. Longer sequences are scarcely needed in practical room acoustic work.

Periodic unit impulses

Periodic unit impulses are a somewhat rhetorical stimulus, as they have very little use in acoustical measurements. The method is suited for testing electrical impulse responses. It has been included for comparison with the MLS method.

A practical use of periodic unit impulses is the comparison of sound card loopback responses. If the impulse responses acquired by the MLS and unit impulse methods differ largely from one another, the system may exhibit nonlinearities causing severe noise degradation in the MLS deconvolution process. An example of such behavior was found in a laptop PC sound card, which produced excessively poor signal-to-noise results with any length of MLS, but yielded credible loopback impulse responses with periodic impulse excitation. This phenomenon was most likely caused by glitches in the input or output buffering, as even a few randomly lost samples will cause an MLS measurement to fall out of synchrony, resulting in all of the remaining MLS signal energy to be misinterpreted as noise.

Arbitrary stimuli

Any arbitrary stimulus may be used for impulse response measurements. The Matlab environment is well suited for the generation of virtually any reproducible signals, enabling the use of custom signals adapted to the nature of the system under test. The spectral energy distribution of the stimulus naturally affects achievable signal-to-noise ratios at different frequency bands.

The deconvolution operation performed by FFT/IFFT post-processing also sets the practical limitations of limited signal length or periodicity.

A list of the corresponding Matlab functions can be found in section 4.3.7.

Data acquisition

The actual measurement operation is performed by an external executable, which accesses the sound hardware on the system. The stimulus and response data are exchanged between Matlab and the sound card handler as multichannel wave files, and the measurement settings are passed in a text file. See the last section in this chapter for details on the `recplay.exe` sound card handler.

A single copy of the stimulus data is stored in a wave file from Matlab. Multiple stimuli are supported by using a multichannel wave file.

Parts of the measurement setup relevant to the output and acquisition process are stored in a text file, which is accessed by the sound card interface application. This contains parameters such as the required sample rate, input and output devices and the number of stimulus cycles. A full description of the file is given in Appendix C.

Matlab calls the interface application which executes and returns after storing the acquired responses in another wave file. The responses are then read to the Matlab workspace.

In order to enhance performance, custom MEX functions were written for reading and writing wave audio files in Matlab, replacing the internal `wavread` and `wavwrite` functions. The use of compiled code resulted in a major increase in speed for handling large wave audio files.

Input peak level detection

In order to ensure reliable results, each channel of the input data is searched through for peak levels, which are displayed on the Matlab console. If any full-scale (*i.e.* 0 dB FS) sample values are found, a warning is issued on possible clipping of the input signal.

Pre-averaging

If pre-averaging is enabled, the response data cycles are averaged to form a single response for each channel. The measurement setup contains options for excluding the first stimulus cycle, which is commonly needed for ensuring that the system response has settled in MLS or other measurements using a continuous stimulus. The last cycle may also be excluded, if there is reason to suspect that the input and output signals may be offset from each other, causing the stimulus to fall before the end of the last response cycle. This effect is possible with low-cost sound cards, which may exhibit a random delay between the input and output signals.

If pre-averaging is disabled, the system extracts a single response cycle, from which the impulse response is evaluated.

As the wide range of algorithms and parameters used for post-averaging are application specific, no implementation has been made in the current measurement software.

4.3.2 Calculation of impulse responses

The current stimulus type is reflected in the methods applied by the measurement system for calculating the impulse response.

MLS stimulus

For MLS stimuli, the impulse response is calculated in a Matlab MEX function using the Fast Hadamard Transform method as described in Chapter 2. The MEX function takes the MLS response and two permutation vectors, which are generated along with the maximum length sequence. The function returns the periodic impulse response.

Unit impulse stimulus

For a periodic unit impulse stimulus, no further post-processing is required.

Arbitrary stimulus

For an arbitrary stimulus, the impulse response is calculated by FFT deconvolution. First, the stimulus and response signals are Fourier transformed. The response data is then divided by the stimulus data in the frequency domain, resulting in the complex frequency response function of the system under test. Finally, this result is returned to time domain using the inverse Fourier transform, resulting in the impulse response function.

The deconvolution method described above may be further utilized for any measured impulse responses in order to exclude the unideal system loopback response from the result. The stimulus signal for the deconvolution is replaced by the loopback response of the sound card or measurement system. For stable time-invariant systems, this results in a very accurate impulse response of the system under test.

A list of the corresponding Matlab functions can be found in section 4.3.7.

4.3.3 Conditioning of impulse responses

Response compensation

Measured impulse responses may be compensated by arbitrary reference signals with FFT deconvolution. The system loopback response may be used as a reference to compensate for unidealities in the signal chain. Another possibility is to use the compensation for inverse filtering together with filtered responses.

Response truncation

Resulting impulse responses may optionally be truncated to a predetermined length. This can be used for removing an excessive noise tail, if long responses with high signal-to-noise ratios are applied for measuring noisy systems with a short decay.

4.3.4 The analysis process

The resulting impulse responses form the basis for all further analysis in the measurement system.

The analysis process is outlined in figure 4.6.

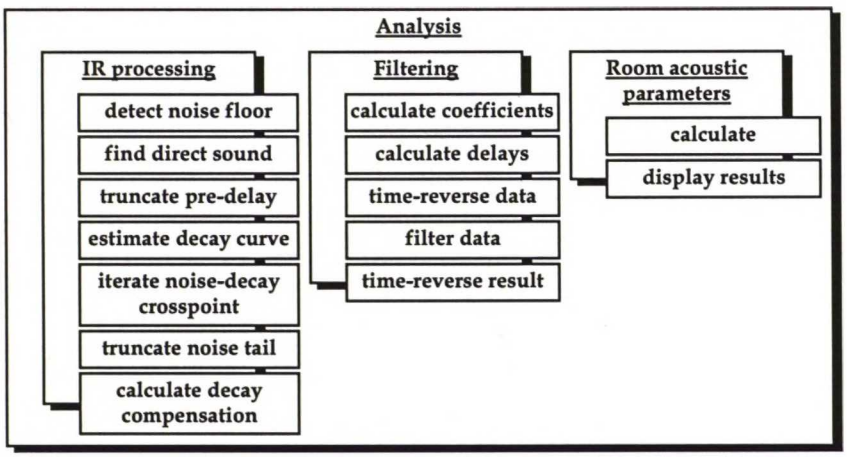


Figure 4.6. An overview of the analysis process.

Processing of impulse responses for analysis

The Lundeby method [Lundeby *et. al.* 1995] for determining the decay/noise knee point and the compensation energy has been implemented in the IRMA system. The implementation is described in figure 4.7 on a functional level.

A list of the corresponding Matlab functions can be found in section 4.3.7.

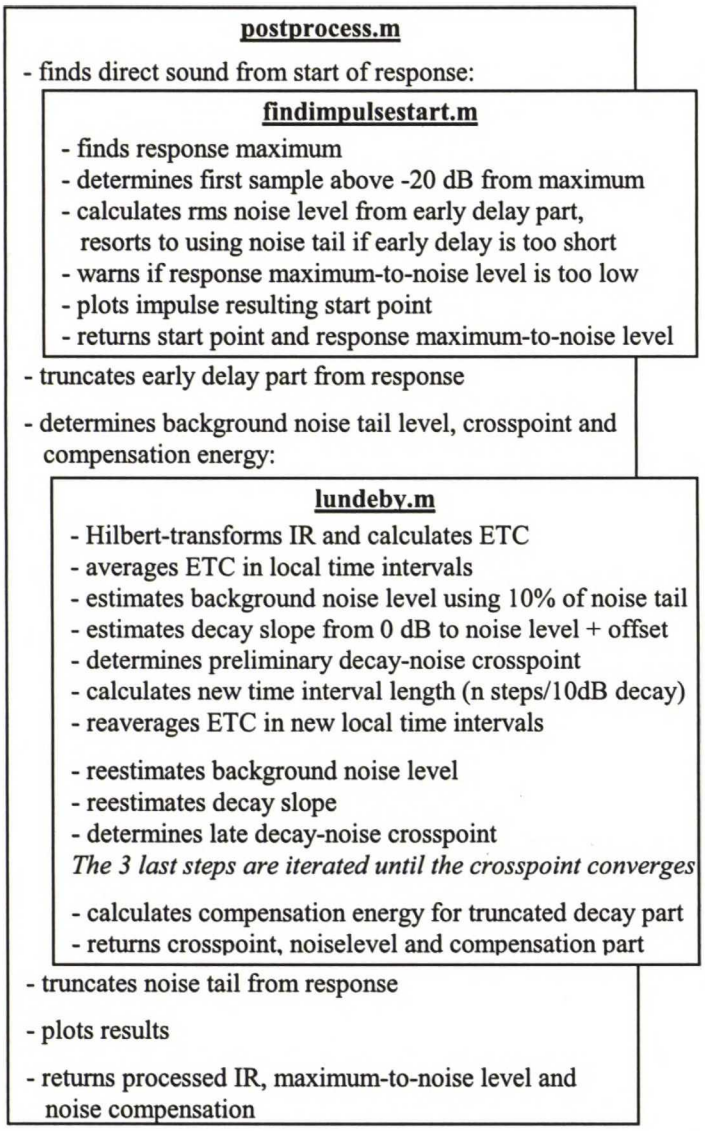


Figure 4.7. Response postprocessing for T_{60} calculation.

Filtering

A set of functions was included to filter the measured impulse responses in standard octave and third octave bands. The filters were realized using IIR implementations of standard analog octave and third octave filters, as defined in the IEC 1260 standard [IEC 1260, 1995].

The filter coefficients are calculated using functions of the Octave toolkit [Couvreur 1997]. The filtering process is described in figure 4.8 on a functional level. A list of the corresponding Matlab functions can be found in section 4.3.7.

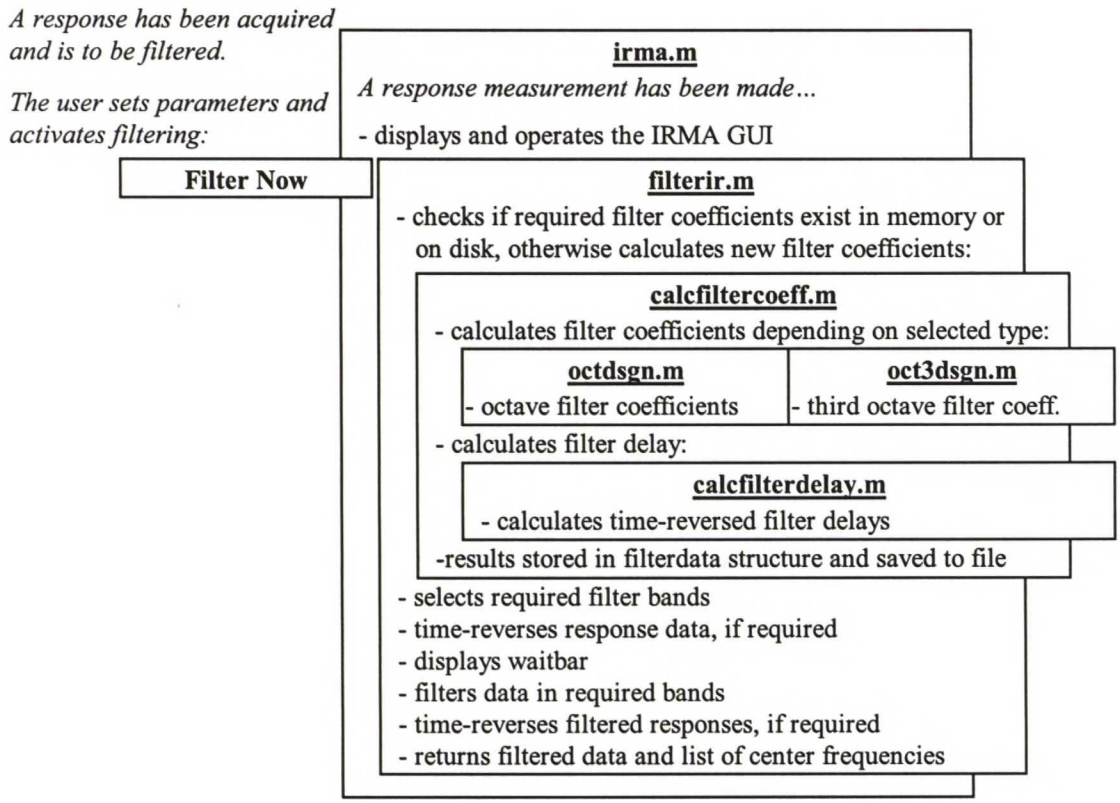


Figure 4.8. The filtering process.

Calculation of room acoustic parameters

The software includes a set of functions for calculating room acoustic parameters as defined in the ISO 3382 standard. The supported functions are listed in Table 4.1 below.

Table 4.1. Functions for calculating room acoustic parameters.

parameter	function	description
T_{60}, T_{10}	schroeder.m	Values for T_{10}, T_{20}, T_{30}, and T_{user} and their correlations. Implemented using the Schroeder integration method. Plots for T_{60} / frequency band as well as raw Schroeder plots.
C_{50}, C_{80}, D_{50}	c50.m, c80.m, d50.m	Implemented
LEF, LF	lef.m, lf.m	Implemented
IACC, IACF	iacc.m	Implemented

A list of the corresponding Matlab functions can be found in section 4.3.7.

4.3.5 The IRMA graphical user interface

A graphical user interface was designed for easier operation of the system for making measurements. As viewing and editing of the measurement setup structure is inherently difficult on the Matlab command line, the GUI provides simple access to the primary measurement, system and hardware setup parameters defined in the setup structure. The GUI also serves as a front end for performing IR measurements and basic analysis. This makes using the measurement application far more obvious than the various functions and scripts involved beneath.

The IRMA GUI window consist of a number of dialogs, which are accessible with the button bar at the top of the window. The dialogs are divided by function, and are described in the following.

The GUI elements and their related measurement setup parameters are also described in the measurement setup documentation, see Appendix C.

Stimulus

The Stimulus dialog offers a choice of stimulus type along with stimulus parameters, the sample rate and the number of stimulus cycles. The current settings are portrayed in the cycle and stimulus length fields, which help to select stimuli of suitable length for the system under test.

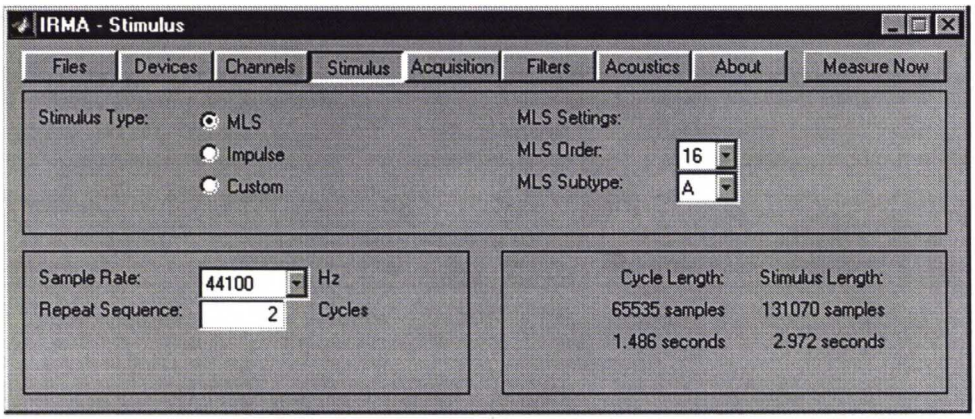


Figure 4.9. The IRMA Stimulus dialog.

The available stimuli are the MLS, unit impulses, as well as any arbitrary stimuli which may be generated or imported into the Matlab workspace.

MLS settings include sequence order and subtype, which corresponds to the location of the feedback taps in the MLS generating register. Sequences from order 2 to order 24 are available, although the practical range is likely to be from order 12 to 20.

Unit impulses are included for simple comparison of electrical loopback responses with those acquired by the MLS method.

The measurement system is not fixed to the MLS method, as any arbitrary "custom" stimuli may be used at will. Impulse responses for custom stimuli are calculated using FFT deconvolution.

The available sample rates for the current sound hardware are 11025, 22050, 44100 and 48000 Hz.

Acquisition

The Acquisition dialog contains settings for response pre-averaging, as well as for response compensation, truncation and storage.

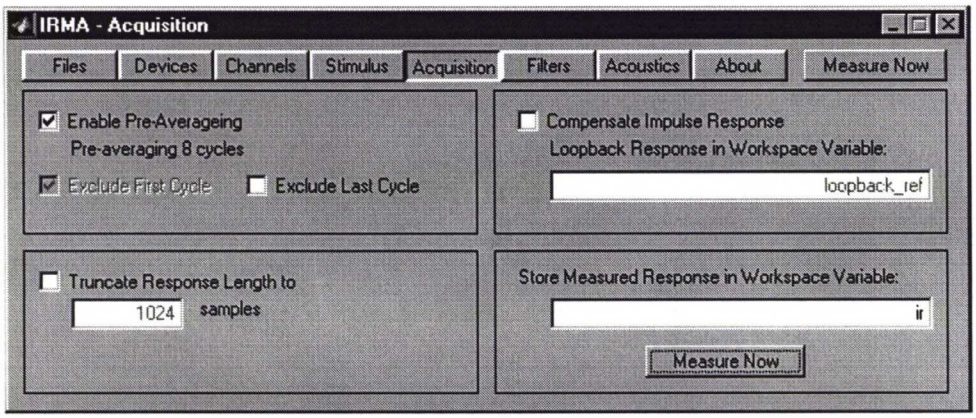


Figure 4.10. The IRMA Acquisition dialog.

Pre-averaging can be used together with the repeat cycles setting to pre-average a number of raw response cycles prior to calculating the impulse response. The first and last stimulus cycles may be excluded from the average. In MLS measurements, the first cycle is always discarded. The last cycle may also be discarded, if the sound hardware exhibits synchronization problems, so that the stimulus ends before acquisition is finished.

The *Measure Now* button is used for performing a measurement. The resulting impulse responses are stored in the Matlab workspace.

Impulse responses may be compensated with other responses. This can be used for compensating unideal system loopback responses. Another possible application is inverse filtering, when used in conjunction with a filtered custom stimuli, such as pink filtered MLS.

Responses may also be automatically truncated to desired length after measurement, enabling the removal of long noise tails.

Filtering

The GUI offers a front end for filtering the measured impulse response in a number of ISO standard octave and third octave bands for the calculation of room acoustic parameters.

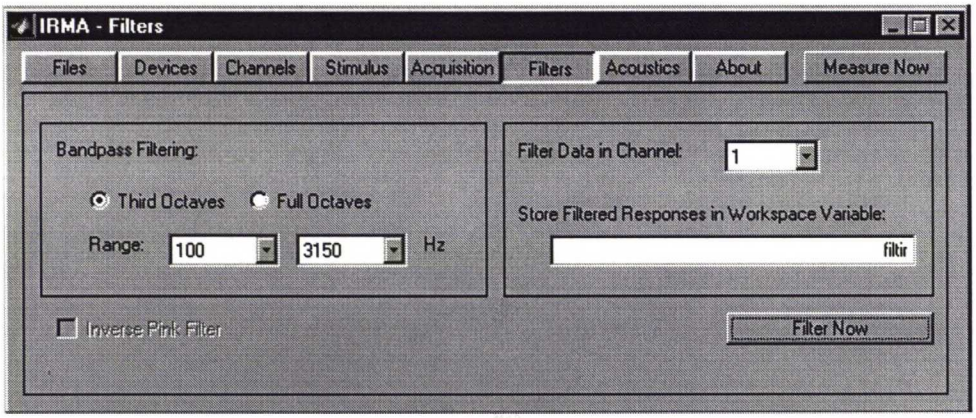


Figure 4.11. The IRMA Filters dialog.

Measurement channels

The channel setup dialog is used to select the input and output devices for a measurement. Each device consists of two-channels. Monophonic stimuli are available at both channels of an output device.

The channel pair approach is due to the way that Windows enumerates system sound devices as independent stereo devices. These can represent either actual two-channel stereo sound cards, or logical two-channel devices that are subdivisions of multichannel hardware.

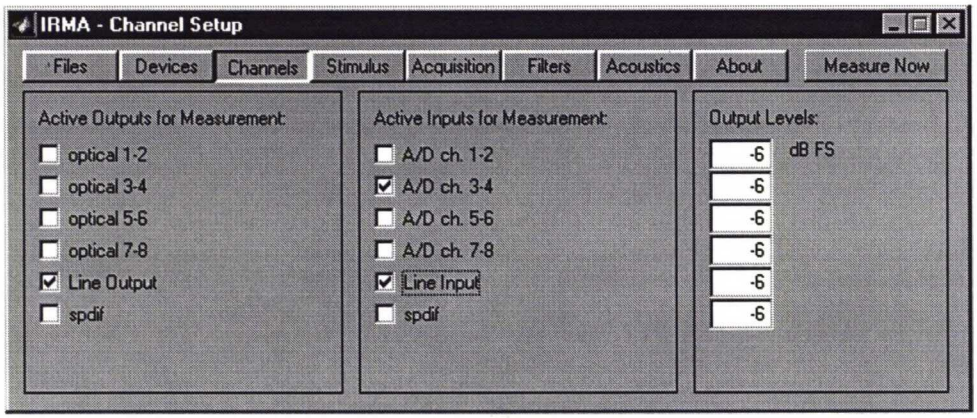


Figure 4.12. The IRMA Channel Setup dialog.

The stimulus signal levels may be controlled for each output device. The level is set in the signal before writing it into the stimulus wave file for the recplay application. Sound hardware output levels must be set by the user by adjusting the output gains on a sound card mixer application.

All input levels are set by the user by adjusting the input gains on a sound card mixer application or an external A/D converter.

System audio devices

The system input and output devices may be labeled by the user according to connection. This can help keep track of multichannel measurement channels.

The hardware device names are contained in the setup structure, and may be edited to match specific hardware setups.

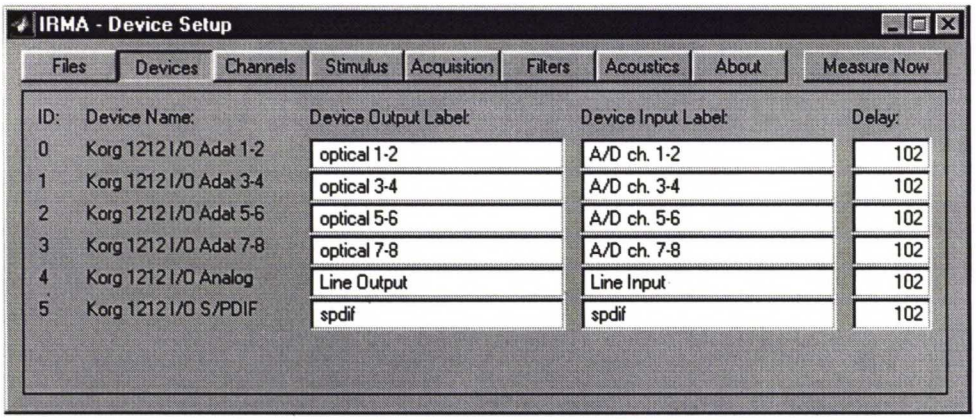


Figure 4.13. The IRMA Device Setup dialog.

Loopback delays may be compensated for each sound device in order to syncronize response starting points. The compensation is performed by acquiring a number of extraneous samples before the beginning of the first response cycle. Alternative delay parameters are included in the setup structure for systems that exhibit negative delay, *i.e.* recording starts ahead of playback.

File management

The Files dialog provides the means for naming, storing and recalling measurement setup structures for different applications.

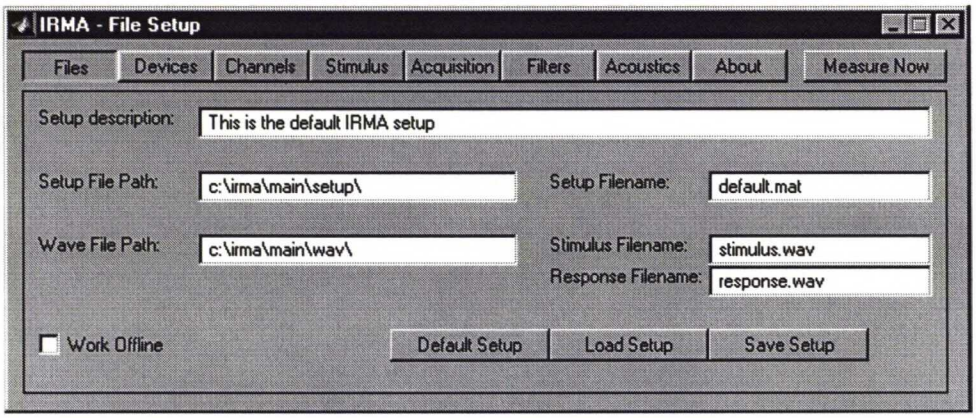


Figure 4.14. The IRMA File Setup dialog.

4.3.6 The IRMA measurement setup structure

A global measurement setup structure was defined for storing the large number of parameters incorporated with the measurement process.

Along with the main application components, most of the included auxiliary functions also access and share the setup data. This aids manual analysis by reducing the amount of required function input parameters, and reduces errors

by constraining the measurement settings to a single location. Together with the IRMA GUI, the setup structure binds the functions together into an application.

A complete documentation of the setup structure is included in Appendix C.

4.3.7 Implemented Matlab functions

A large number of custom Matlab functions were written for the measurement application. Although a large part of the functions can be directly accessed by the IRMA GUI, nearly all functions are also available for common Matlab use.

The functions can be roughly categorized into several types: stimulus generation, GUI and measurement handling, IR calculation and post processing, filtering, room acoustic parameters, data plotting and output, and a group of miscellaneous utilities for handling the measurement setup and different file types.

GUI and measurement

Table 4.2. GUI and measurement functions.

`irma.m`	graphical user interface for the measurement application
`measureir.m`	performs IR measurement
`compensateir.m`	compensates IR using deconvolution
`makestimulus.m`	generates stimuli
`makeirsetup.m`	script for creating the default measurement setup
`recplayini.m`	writes measurement parameters to file for RECPLAY.EXE

MLS

Table 4.3. MLS related functions.

`mls.m` (`mls.dll`)	generates maximum length sequence
`mlsxcorr.m` (`mlsxcorr.dll`)	calculates impulse response using the Fast Hadamard Transform
`mlsfilecreate.m`	generates and stores MLS sequences in .mat and .wav files

IR processing

Table 4.4. IR processing functions.

postprocess.m	post processes impulse response for calculation of room acoustic parameters
findimpulsestart.m	determines start of direct sound and early delay maximum-to-noise level
levelsearch.m	detects first occurrence of given level
lundeby.m	determines decay-noise crosspoint, maximum-to-noise level and noise compensation
timeaverage.m	time-averages a squared impulse response into intervals
estimatenoiselevel.m	estimates background noise level of impulse response
estimateslope.m	estimates slope of decay from level1 to level2

Filtering

Table 4.5. Filtering functions.

filterir.m	filters data in octaves/third octaves
calcfiltercoeff.m	calculates filter coefficients
calcfilterdelay.m	calculates filter delays
octave\octdsgn.m	calculates ISO octave filter coefficients
octave\oct3dsgn.m	calculates ISO third octave filter coefficients

Room acoustic parameters

Table 4.6. Room acoustic parameters.

rtcalc.m	calculates the reverberation times for a set of filtered IR bands
schroeder	calculates the reverberation times for an IR using Schroeder integration

Table 4.7. Energy parameters.

c50.m	calculates C_{50}
c80.m	calculates C_{80}
d50.m	calculates D_{50}
centertime.m	calculates T_S
energy.m	calculates signal energy in specified time interval

Table 4.8. Binaural parameters.

iacc.m	calculates inter-aural cross-correlation from an IR pair
iacccalc.m	calculates IACC values for two sets of filtered IR bands
lf.m	calculates Early Lateral Energy parameter LF ($\cos^2$ method)
lfc.m	calculates Early Lateral Energy parameter LFC (cos approximation)

Miscellaneous functions

Table 4.9. Output and plotting.

`plotir.m`	creates amplitude-time and energy-time plots
`plotfft.m`	creates frequency-time plots

Table 4.10. File I/O functions.

`wavread.dll`	reads wave files, MEX version for enhanced speed
`wavwrite.dll`	writes wave files, MEX version for enhanced speed
`timread.m`	reads MLSSA time domain data from .tim files
`timwrite.m`	writes MLSSA time domain data to .tim files
`frqread.m`	reads MLSSA frequency domain data from .frq files

Table 4.11. Accessory functions.

`ca.m`	calculates and displays room acoustic parameters from filtered IR
`schroederseq.m`	generates Schroeder sequences
`checkglobal.m`	checks whether global variable resides in memory and is nonempty
`distdata.m`	generates discrete points in distance for data at sample rate f_s
`freqdata.m`	generates discrete points in frequency for data at sample rate f_s
`timedata.m`	generates discrete points in time for data at sample rate fs
`linreg.m`	linear regression and correlation
`cconv.m`	circular convolution
`cxcorr.m`	circular crosscorrelation

4.4 The IRMA sound card interface software

Although the Matlab environment is well suited for the creation, analysis, display and storage of measurement signals and data, it does not offer direct methods for data acquisition†. In order to realize the measurement system, a sound card interface was required. A sound card handler was implemented as a custom 32 bit Windows application for accessing Windows sound drivers directly from the Matlab environment.

Functional Requirements

Essentially, the sound card interface should be capable of simultaneous multichannel playback and recording, using arbitrary sound devices on the system. In addition, a transfer method is required for handling the exchange of data with Matlab. A module capable of fulfilling these tasks would inherently be applicable to any measurement method or stimulus.

The requirements for data transfer consist not only of an audio data block format suitable for passing the stimuli and responses between the applications, but also of a number of measurement settings passed from Matlab to the sound card interface. These parameters include choice of playback and recording devices, pointers or filenames to stimulus and response data, information on data format and sample rate, the number of cycles to repeat, and possible settings for pre-averaging the responses.

Methods for Realization

Choosing the right tools for programming the sound card interface required some insight into the application program interfaces available in the Windows and Matlab environments.

The Microsoft Windows operating systems (Windows 3.x/95/98/NT) offer a common application program interface for handling sound cards. The Windows Sound API forms an interface for any sound cards that are installed on the system with Windows device drivers. The Sound API contains functions for querying the number and properties of sound devices on the system, as well as low-level functions for reading, writing and handling audio on the sound devices.

Sound cards are available under Windows as enumerated stereo sound devices. Windows sets a unique device ID for each sound device with integers starting

† During the course of coding the measurement system, The Mathworks released the Data Acquisition Toolbox, a new product which enables direct data acquisition in the Matlab environment by using a Windows sound card and other common types of data acquisition hardware. The Data Acquisition Toolbox offers much of the functionality required for the purposes of this project. However, the latency times between input and output channels are large, and are noted to exhibit some variance between subsequent measurements.

from 0. This ID is used by the Sound API commands for specifying the target device. For ordinary stereo sound cards, each sound device relates to an independent piece of hardware. Multichannel cards, however, are available as several virtual stereo devices. From the Windows viewpoint, sound devices are independent of one another, and each device can be accessed separately with unique settings. For multichannel sound cards, the individual virtual sound devices are usually constrained to common settings, although they can be accessed individually. Multichannel sound cards also operate the individual channels from a single clock source, which results in proper inter-channel synchronization. [Microsoft 1997]

Matlab version 5 includes an API for interfacing with programs and functions written in C or Fortran. The Matlab MEX compiler enables these to be built into custom dynamic link libraries, which execute as MEX functions under the Matlab environment.

The exchange of data between MEX functions and the Matlab environment is performed directly in memory, using pointers to the arguments passed with the Matlab workspace. The exchanged data must be in standard Matlab data formats, such as matrices of type double float. If the MEX function uses other data types internally, it must be ensured, that all data exchanged with Matlab is first converted from one type to another. For functions with a large number of passed arguments, this requires a considerable amount of supplementary code for data conversion. Conversion of large data sets from one data type to another may also result in an overhead in memory allocation and processing times.

It is also possible to launch separate executable programs from the Matlab environment, and pass parameters to the programs as command line arguments. There is no direct method of interaction from the programs to Matlab, so disk storage must be used for most data transfers. However, it is relatively simple to first store the necessary data into a disk file from Matlab, then execute the external program passing optional arguments, and finally read back any results from disk after the execution of the external program has finished.

This results in two alternative realizations of the sound card interface. The choice of data transfer method between the interface and Matlab environment could be implemented either by an external interface application using disk storage for data transfer, or by a Matlab MEX function interfacing directly with the Matlab workspace. The internal functionality of these alternatives is largely similar, except for the data transfer methods.

The realization of an external executable is less constrained than the Matlab MEX format [Mathworks 1998]. Both types yet require the definition of data exchange formats between Matlab and the interface application.

The external interface requires a common disk file format for storing the audio data. The RIFF wave format is a standard file format capable of storing both multichannel data and audio format information. The Windows Multimedia API and Matlab both include functions for handling wave audio files, making the format well suitable for the application.

Implementation

The simplest solution for an external interface could be the use of any Windows multichannel recording application, capable of full duplex playback and recording, with wave files for the transfer of audio data. This scheme was actually used for testing early in the project. The simplicity of this method is only hypothetical, however, as there is no way of controlling the setup parameters and measurements directly from Matlab. Instead, any measurement setup parameters must be set manually in the recording application, and the data must be saved and loaded separately for each measurement both in Matlab and in the recording application.

In order to initiate and control the measurement directly from Matlab, a custom interface is required. An independent executable could be implemented by storing the stimulus and response data in wave files, and storing the setup parameters in a text file. Although this method requires disk access, an independent sound card interface application seemed more attractive of the two: measurements could be performed and stored in disk wave files, regardless of whether Matlab was being used in conjuction with it. This was well in line with the basic objective of modularity and independence sought in the formulation of the project.

The sound card interface was coded in standard ANSI C, using Microsoft Sound API library functions for interfacing with the sound devices and handling wave audio files. The application was built using the Microsoft Visual C 5.0 compiler as an independent 32 bit Windows executable.

The interface is coded in a modular fashion, with separate function sets for sound card access, wave file access, and the handling of measurement settings. The code can be converted into a Matlab MEX function by adding functions for the handling and conversion of Matlab data formats. For heavy multichannel use, some optimization of memory allocations related to the conversions may also be necessary. The sound card control application can also be used as a basis for further stand-alone implementation of arbitrary multichannel audio measurement software.

The sound card interface for the IRMA measurement system includes the following files:

Table 4.12. Files related to the IRMA sound card interface application.

`recplay.exe`	the Win32 command line measurement application
`recplay.ini`	text file containing acquisition settings
`stimulus.wav`	single, stereo or multichannel stimulus in wave file
`response.wav`	stereo or multichannel responses in wave file

5 Evaluation of the implemented system

This chapter begins with a series of comparative measurements to evaluate the IRMA system's functionality, and to verify the validity of the resulting impulse responses and room acoustical criteria. This is followed by a group of simulations to evaluate the effects of the bandpass filters used in data analysis. The chapter ends with a discussion on the effects of the MLS measurement method and the various response processing methods.

5.1 Room acoustic measurements

A series of comparative measurements was carried out to evaluate the performance of the IRMA system and to compare the resulting impulse responses and room acoustic criteria with results acquired with an acknowledged measurement system. A MLSSA analyzer [Rife 1996] was used as the reference system.

The measurements were performed by successively measuring the same source–receiver pairs in two different acoustical spaces with both systems. Between measurements, no changes were made to the acoustical system or the measurement setup, except for interchanging the MLSSA system with the IRMA system. The measurement hardware used is listed in the table below.

Table 5.1. Measurement hardware.

measurement systems	IRMA, DRA Labs MLSSA
sound source	meeting room: Genelec 1029A active loudspeaker concert hall: HUT Acoustics Laboratory's omnidirectional dodecahedral source
receiver	B&K 4006 omnidirectional microphone
microphone amplifier	Sonolab custom preamp

Only a single source-receiver pair response was included for each measurement, as the main objective was to compare the results from the two measurement systems. Multiple source-receiver pairs would naturally have been measured and averaged, had the actual acoustical parameters of the room been investigated.

Standard room acoustical criteria were evaluated from the acquired impulse responses on both systems. To enhance the comparison, both automatic and manual evaluation of parameters was performed on the MLSSA system. In addition, the impulse responses acquired with MLSSA were imported into Matlab and analyzed with the IRMA system. This makes it possible to compare the systems' response acquisition and analysis capabilities independently.

The measurement results and related discussion are found in chapter 6.

5.1.1 Meeting room

A basic verification of the IRMA system was carried out by measuring the responses of a source-receiver pair situated in an unoccupied office meeting room. The source and receiver were placed in the room in accordance with building acoustical criteria for reverberation measurement.

Table 5.2. Meeting room measurement setup parameters.

parameter	IRMA	MLSSA
sample rate, Hz	48 000	60 610
measurement bandwidth, Hz	20 000	20 000
MLS order	16	16
pre-average cycles	8	8

5.1.2 Concert hall

Another comparison test was carried out at the concert hall of the new Helsinki Conservatory building. The measurements studied are extracts from a larger series of acoustical measurements performed at the hall. The example response is taken from the onstage receiver point P2, with the source at point S2. The points were chosen according to the recommendations for source-receiver points in concert hall measurements [Gade 1989].

Table 5.3. Concert hall measurement setup parameters.

parameter	IRMA	MLSSA
sample rate, Hz	44 100	30 075
measurement bandwidth, Hz	20 000	10 000
MLS order	17	16
response length,s	2.972	2.179
pre-average cycles	4	8

5.2 Filter simulations

Band-pass filtering forms an important part in the analysis of room acoustic impulse responses prior to the calculation of room acoustic criteria. In this section, the properties of the IRMA system's filter realization are examined.

The IRMA system uses the digital IIR counterparts of standard analog octave and third octave filters. In section 3.2.3, general aspects of these filters were discussed, along with their digital realization.

This section examines the decay times and delays produced by the IRMA system's filter implementation for both forward and time-reversed methods. In addition, a brief mention is given to other limitations incorporated with the IIR filter implementation.

5.2.1 Filter decay and center times

Filter decay and center times were calculated from impulse responses of the octave and third octave bandpass filters that are used in analysis. The filter

decay times limit the shortest decays measureable by the system, whereas the center times show the magnitude of delay occuring in the filtering.

The results, calculated for both time-reversed and ordinary octave and third octave bandpass filters at a sample rate of 44.1 kHz, are shown below in both numerical form and as Schroeder-integrated decay curves.

Table 5.4. Schroeder integrated decay times for time-reversed and ordinary octave bandpass filters with a sample rate of 44.1 kHz. Results of linear regression from 0 dB to –60 dB together with correlation coefficients. Filter response center times shown for comparison.

f_c	125	250	500	1000	2000	4000	8000
ordinary, T_{60} (ms)	71.31	35.76	17.91	8.97	4.47	2.2	0.99
r	-1	-1	-1	-1	-1	-1	-1
time-reversed, T_{60} (ms)	14.05	6.74	3.41	1.62	0.44	0.18	0.06
r	-0.85	-0.83	-0.84	-0.83	-0.62	-0.65	-0.73
center time, T_S (ms)	8.42	4.22	2.12	1.06	0.53	0.27	0.14

Table 5.5. Schroeder integrated decay times for time-reversed and ordinary third octave bandpass filters with a sample rate of 44.1 kHz. Results of linear regression from 0 dB to –60 dB together with correlation coefficients. Filter response center times shown for comparison.

f_c (Hz)	ordinary, T_{60} (ms)	r	time-reversed, T_{60} (ms)	r	center time, T_S (ms)
100	214.88	-1.00	45.25	-0.92	32.39
125	170.88	-1.00	35.81	-0.92	25.73
160	135.24	-1.00	28.43	-0.91	20.43
200	107.62	-1.00	22.6	-0.92	16.23
250	85.49	-1.00	17.86	-0.91	12.89
315	67.91	-1.00	14.19	-0.91	10.24
400	53.94	-1.00	11.17	-0.91	8.14
500	42.84	-1.00	8.88	-0.91	6.46
630	34.03	-1.00	7.03	-0.91	5.13
800	27.02	-1.00	5.52	-0.90	4.08
1000	21.46	-1.00	4.46	-0.91	3.24
1250	17.04	-1.00	3.45	-0.90	2.57
1600	13.52	-1.00	2.77	-0.91	2.04
2000	10.73	-1.00	2.06	-0.88	1.62
2500	8.51	-1.00	1.64	-0.89	1.29
3150	6.75	-1.00	1.34	-0.90	1.02
4000	5.34	-1.00	0.72	-0.68	0.81
5000	4.2	-0.99	0.53	-0.67	0.65
6300	3.32	-0.99	0.4	-0.67	0.51
8000	2.61	-0.99	0.3	-0.67	0.41
10000	2.04	-0.99	0.21	-0.68	0.33

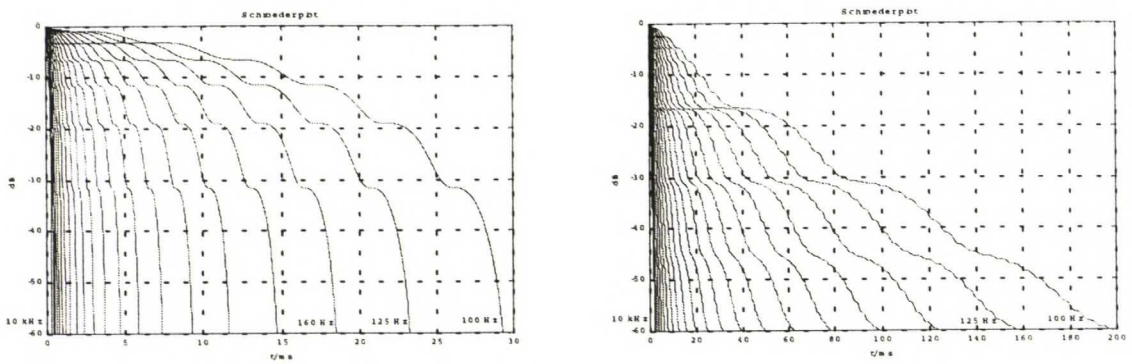


Figure 5.1. Schroeder integrated decay times for a) time-reversed and b) ordinary third octave band filters. fs = 44100 Hz. Note the different time scales between plots.

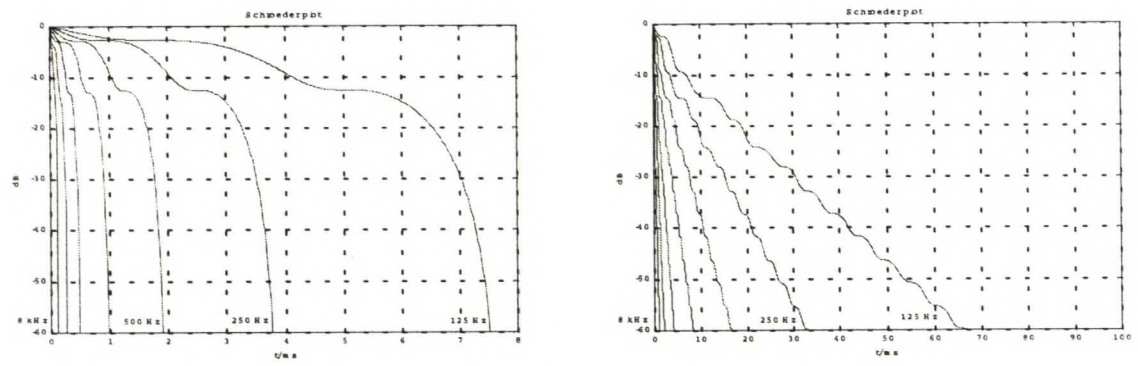
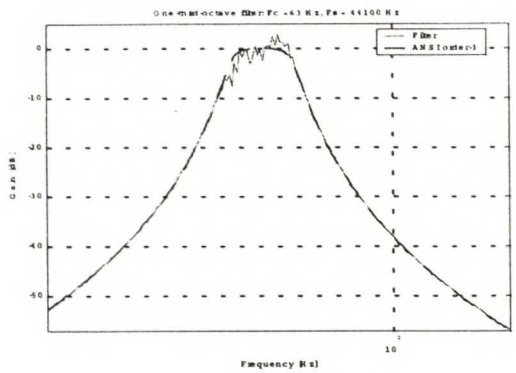


Figure 5.2. Decay times for time-reversed and ordinary octave band filters. Note the different time scales between plots.

5.2.2 Limitations caused by filter implementation

The IIR butterworth octave and third-octave filters are found to be applicable only inside the frequency band in the range of $f_s/500$ to $f_s/4$, where f_s is the sample rate. This corresponds to the range of 100 Hz to 10 kHz for a sample rate of 44.1 kHz. Beyond these limits, the filter shapes fail to fill the standard requirements [IEC-1260:1995].

Examples of discrepancies in the bandpass filter implementation at low frequencies are shown in figure 5.3a. These are due to effects of numerical inaccuracy in the digital filter coefficients. High frequency inaccuracies are shown in figure 5.3b. Here, the the Nyquist frequency acts as a limiting factor to the filter's upper side lobe.



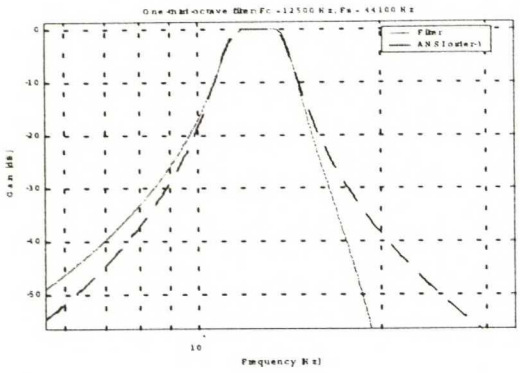


Figure 5.3a. Limitations in IIR filter realization at low frequencies.
Figure 5.3b. Limitations in IIR filter realization at high frequencies.

5.3 Effects of the MLS measurement method

MLS drawbacks

The main drawbacks of the MLS method are the rather strict requirements for system linearity and time invariance. These can result in serious pitfalls, as severely misleading results can be acquired from non-LTI systems, if the unideal conditions are not suspected. In practice, MLS measurements should only be carried out under well known stable condition, unless the user has a reasonable amount of expertise along with the possibility to check for errors by carrying out some analysis of the measured responses in-situ.

As with all methods based on deterministic stimuli, MLS measurements cannot be used with systems that do not accept an external stimulus. In such cases, ordinary 2 channel FFT methods are commonly applied.

Although the MLS method can provide excellent signal-to-noise ratios under ideal conditions, the background noise levels acquired in practical measurements are often found to fall far from this ideal. Some of the reasons for this are discussed in the following.

Dynamic Range

Decay range is the level difference in decibels between the initial maximum of the decay curve and the point where the decay approaches the noise floor [Bradley 1995].

The decay range may be increased by increasing the stimulus level and by decreasing the ambient noise level. Averaging of several measurements forms another method for increasing the measurement signal-to-noise ratio. With the MLS method, the use of longer sequences also effectively increases the measurement SNR, as the total energy emitted by the source is increased with time. The cross-correlation method involved also effectively rejects uncorrelated noise from the measurement.

However, in practice, a large decay range is not always obtained, and it is essential to optimize the measurement parameters to ensure that an extended decay range is achieved.

Significant limitations to the decay range can be caused by an inadequate SNR between the source and the ambient noise levels, loudspeaker distortion, too short an MLS length, or an inadequate number of averages.

Distortion

Several studies have been made on identifying and quantifying the effects in MLS impulse responses that are caused by various types of distortion [Rife, Vanderkooy 1989], [Dunn, Hawksford 1993], [Dunn, Rife 1994], [Vanderkooy 1994]. High levels of distortion commonly lead to wildly inaccurate results, whereas certain types of weak nonlinearities produce easily identifiable artefacts in the response. Low-order harmonic distortion has been shown to produce a number of echo-like images that are added to certain time-delayed locations in the response. These image locations change, as the MLS generating tap locations or the sequence length are varied. This gives means to identifying the presence of distortion, and possibly even for cleaning up the response by discarding the artefact images. [Dunn, Hawksford 1993]

Distortion must however be seen as a severely limiting factor in MLS measurements. With the presence of distortion, increasing the number of acquired averages can fail to provide any improvement to the resulting decay range. For moderate levels of distortion, a longer sequence length can yet be used to improve the measurement SNR. Thus the best choice of sequence length may be much longer than the length of the decay exhibited by the system, as the available decay range may be increased.

The adverse effects of loudspeaker distortion must be understood. When the level of an already distorted source is further increased, the decay range of the MLS response decreases, whereas decreasing the source level can actually increase the decay range due to reduced distortion effects.

Time-variance

Time variance in the system under test may occur as changes of temperature and wind, or as changes in geometry with the presence of moving objects, subjects or transducers such as a turning (or drooping!) microphone boom. Outdoors variations in temperature and wind may cause phase shifts and changes of sound speed.

The effects of wind and temperature drifts in indoor and outdoor MLS measurements have been studied by [Vorländer, Kob 1996, 1997], with results that are of practical use in building acoustics. The authors derive formulae for error estimates and give maximum permissible values for the variations in temperature in common measurements. For example, in accurate reverberation time measurements, the room temperature should be kept stable within a few tenth of a Kelvin during the total measurement time.

A difficult source of time-variance may be caused by data buffer overruns or underruns in digital measurement systems. Such behavior was found using several low-cost pc sound cards.

In general, MLS responses of time-variant systems exhibit unseemingly high background noise levels. As MLS responses are processed in the time domain, a loss of time synchrony between the system input and output causes the remaining parts of the sequence not to correlate. This can result in large parts

of the stimulus signal to be misjudged as uncorrelated background noise. The presence of time-variance in a system may often be verified by comparing the acquired system response with that produced by a longer sequence. Alternatively, a larger number of averaged cycles may be used instead. Under time-invariant conditions the longer sequence should result in a higher signal-to-noise ratio than the original sequence. If this is not the case, the system most likely exhibits time-invariance.

On a low-quality laptop computer sound card, MLS loopback measurements were found to result in a maximum of 30 dB signal-to-noise for short sequences. Longer responses yielded even worse figures. For comparison, a 60 dB signal-to-noise ratio was measured with the same setup, when using a unit sample sequence instead. Clearly, the sound card exhibited severe defincies in the full-duplex buffering of audio data.

Equalization

Appropriate spectral shaping of the wideband stimulus signal can increase the SNR of the measured response, when more source energy is located at frequency bands with higher levels of background noise. In practice, a pink filter is often used, as it correlates well with the ambient noise spectrum common to halls and rooms. This results in an increased SNR at low frequencies.

The use of equalization to compensate for limitations in the frequency response of a loudspeaker will unfortunately only result in a decreased overall SNR with MLS measurements. At frequency bands requiring high gain, the loudspeaker sensitivity is usually very low, resulting in increased distortion and a lower SNR. On the other hand, the attenuation of accented frequency bands will effectively only decrease the SNR between the source signal and the background noise. In effect, for MLS measurements, the flatness of the source's frequency response should not be a primary aim, as the decay range is generally inversely related to the amount of equalization involved. [Bradley 1996]

As a result, there exists an optimal signal level for MLS measurements, that is affected by both the ambient noise level and the distortion behavior of the sound source. This level varies with frequency, depending on the spectrum of the ambient noise, and the effective frequency range of the sound source.

5.4 Effects of response post-processing methods

5.4.1 Filtering

Alternative filter implementations

In the previous chapter, limitations related to the accuracy of IIR filters near the edges of the sampled bandwidth were shown.

To cover the entire audio frequency range would require the use of FIR filters designed to the standard specifications [IEC 1260, 1995]. At the highest frequency bands, sufficiently steep transition bands would require filters of increasingly large order. An alternative and more efficient approach would be the use of multirate filtering, where the data is first decimated to a lower

sample rate, IIR filtered, and then interpolated back to the original sample rate. However, in order to limit the scope of the thesis work, such functionality was not included. A multirate filtering scheme covering the whole audio band would be relatively simple to implement to the software at a later stage.

Effects of filters on reverberation time

The effects of filter decay time on short reverberation times was discussed in Chapter 3.

Time-reversed filters can also affect EDT values, as they are prone to change the early part of the decay curve. In their recent paper, Kob and Vorländer [Kob, Vorländer 2000] claim, that actually the EDT cannot be determined accurately from time-reverse filtered responses. The T_{20} and T_{30} values are not as prone to error, as they are evaluated over a larger range with the less sensitive –5 dB decay level as their starting point.

Effects of filters on energy-time parameters

The calculation of energy-time parameters involves windowing of impulse responses into early and late time periods. The starting point of the early time period should be located at the arrival of direct sound, and the midpoint dividing the response into early and late parts should be located at a specified time interval (commonly 50 or 80 ms) from the starting point. When band-filtered results are required, however, the effects of filter delays and response energy time smearing are encountered. Filter delay shifts energy in the filtered response forward in time and spreads it over a longer time interval [Bradley, Halliwell 1992].

In order to properly account for these side effects, one of two methods should be employed [ISO 3382, 1997]. In the time-windowing method, the broad-band impulse response is windowed prior to filtering and summing up the energy contained in each part. This assures a static starting point, and prevents response energy from being transferred from the early time period into the late time period [Bradley, Halliwell 1992]. Alternatively, the window correction method can be used to evaluate the starting point and midpoint from each band filtered response in the following manner: The starting point should be located by determining the first point where the filtered signal first rises significantly above the background noise, but is more than 20 dB below the maximum level. The midpoint should then be located by adding the early time interval plus one half of the filter time constant in question to the starting point [Barron 1984]. Although the ISO 3382 standard recommends using the time-windowing method, it has been demonstrated that especially for low frequency bands the window correction method can produce more accurate results [Lundeby 1995].

5.4.2 Processing and analysis of impulse responses

Direct sound delay

Direct sound propagation delay affects not only the calculation of energy-time parameters, but to a small extent also Schroeder integration. With an adequately low background noise level, the pre-delay part of an impulse response contains very little energy compared to the direct sound. This forms a practically non-decaying segment at the beginning of the Schroeder integral,

and acts as an error source for the EDT calculation, where the upper integration limit is taken at 0 dB. The stationary delay segment is thus included in the decay process, overemphasizing the direct sound energy in the decay estimate.

Figure 5.4 portrays the difference by a real example, which yields a difference of about 5 %. The lower correlation coefficient (0.98) also gives a notion of the non-linearity in the decay curve.

Direct sound delay does not affect other reverberation parameters, as their upper limits are taken below 0 dB.

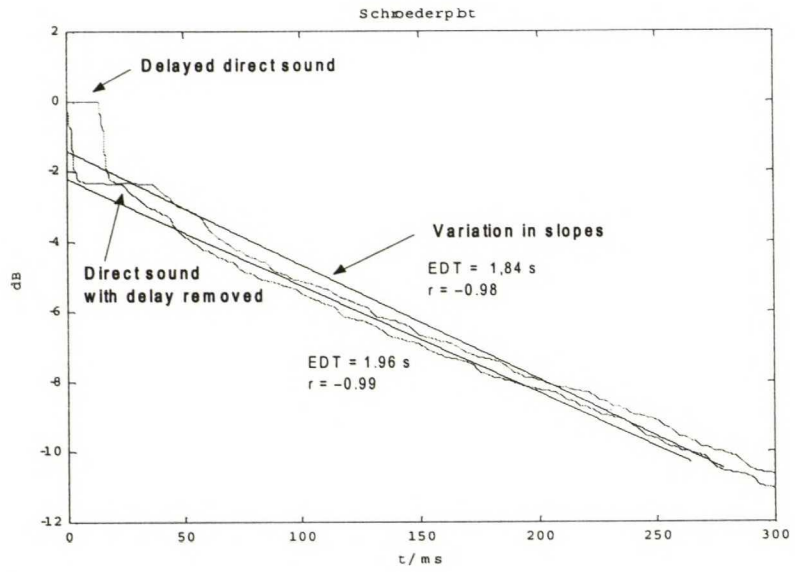


Figure 5.4. The effect of delayed direct sound on EDT calculation.

Noise tail truncation

The crucial element of the Schroeder reverse-time integration is the choice of the integration limit.

Methods for *noise compensation* attempt to improve the accuracy of the decay curve. They are based on an assumption of exponential decay following the upper integration limit, and introduce a correction term containing an estimate of the total energy in the interval from the integration limit to infinity.

A systematic underestimation of T_{30} by not more than 5 % results only, if the integration limit corresponds to a level of the impulse response of at least –45 dB. As a rule of thumb, the impulse response must be evaluated to a level at least 10 dB lower than the value of the lower limit in the evaluation range of the specific reverberation measure. (–20 dB for EDT, –35 dB for T_{20} or –45 dB for T_{30}). [Vorländer, Bietz 1994].

If the background noise is known, the optimum starting point for the reverse-time integration is the point at which the decay line intersects the background noise level.

If the background noise is unknown, special algoritms must be applied to check signal-to-noise ratio and to exclude unreliable results. These algoritms may be tested by comparing the results and the repeatability of the automatic measuring techniques with results obtained by manual evaluation.

The effects of noise tail truncation and noise compensation are shown in Figure 5.5. The upper graph displays the Schroeder integrated impulse response including the background noise tail. In the lower graph, the same response has been truncated at the crosspoint between the decay slope and the stationary background noise. The degrading effects of the noise tail can be clearly seen in the upper graph, which begins to bend toward the noise tail some 10–15 dB above the actual noise level. The lower graph curves much less, enabling more accurate determination of the decay slope.

In addition, the lower curve has been artificially extended by the noise compensation method, in an attempt to extend the late decay slope below the actual noise level. The slight bend at the crosspoint indicates the difficulty of accurately matching the late decay slope.

Instead of ultimately extending the truncation crosspoint to as close as possible to the background noise level, it may in some cases be better to truncate the decay at a higher level that is unaffected by noise, and use the more ideal slope for estimating the noise compensation curve. This method must be used with caution, as double decay may be left unnoticed.

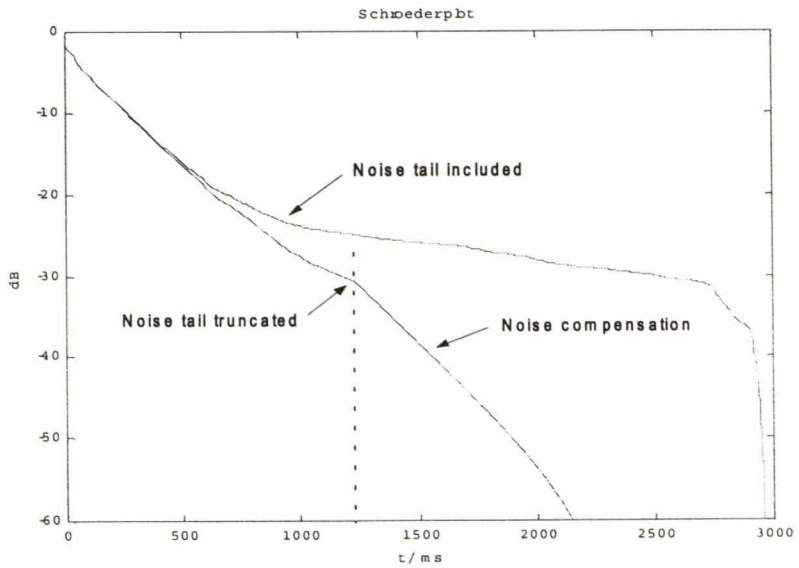


Figure 5.5. Noise tail truncation and noise compensation.

Figure 5.6 shows the same response, with the truncation point now manually shifted about 50 ms backward in time. This reduces the effects of background noise bending, and the noise compensation curve makes a good fit with the actual decay slope.

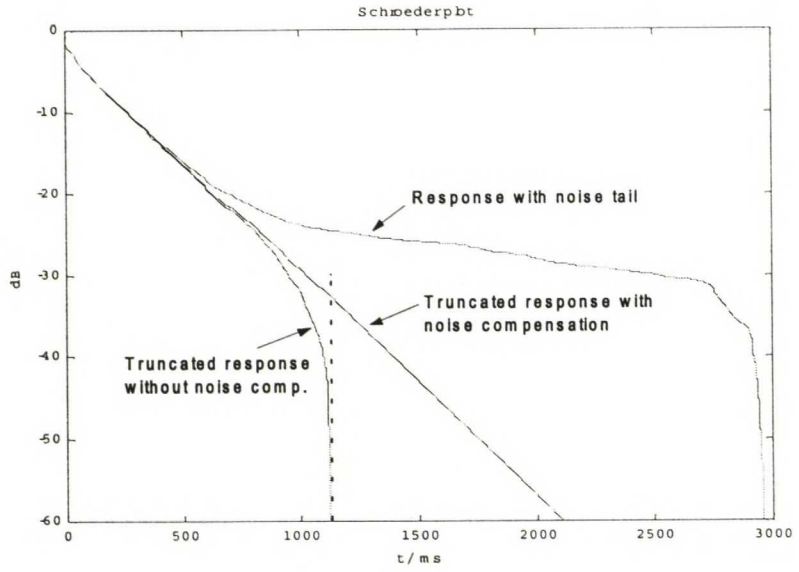


Figure 5.6. Optimal noise truncation with manually adjusted truncation point.

Figure 5.7 shows the effects of the use and lack of noise compensation on the actual reverberation time parameters. The graphs portray the same response truncated at 1200 ms. The upper graph includes noise compensation, whereas the lower graph does not. The results show differences as large as 20 % for T_{30}. It is interesting to note that both calculations yield excellent correlation coefficients regardless of the differences in the actual values. Although correlation can act as a warning for unreliable results, it is seen that this is not always the case. In practice, differences of such magnitude could cause considerable consequences.

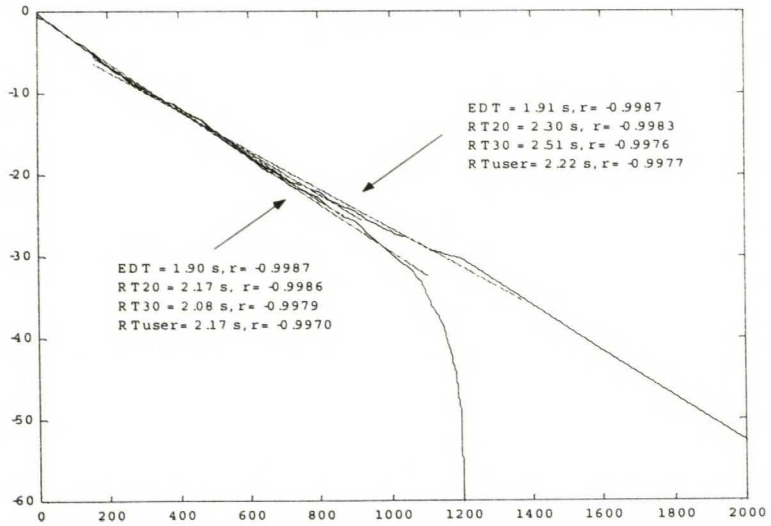


Figure 5.7. Effects of noise compensation on reverberation time parameters.

The Lundeby method

Figure 5.8 displays the operation of the Lundeby method for evaluating the truncation point and the background noise level. The squared response is shown in the background, and two series of time averages of the squared response at front.

The first time average series is used for calculating preliminary values for the background noise level and the full decay slope. A truncation point is located at the intersection of these lines. The next time average series is calculated based on this information, so that the decay part now contains about 5 averages per 10 dB of decay. This new series is used for reestimating the background noise level and the late decay slope. New truncation points are calculated iteratively, until a stable value is reached. The method is outlined in [Lundeby *et al.* 1995], and realized in the `lundeby.m` Matlab function.

The late decay slope and the truncation point are used for calculating an estimate of the remaining decay energy beyond the truncation point, ie. the noise compensation.

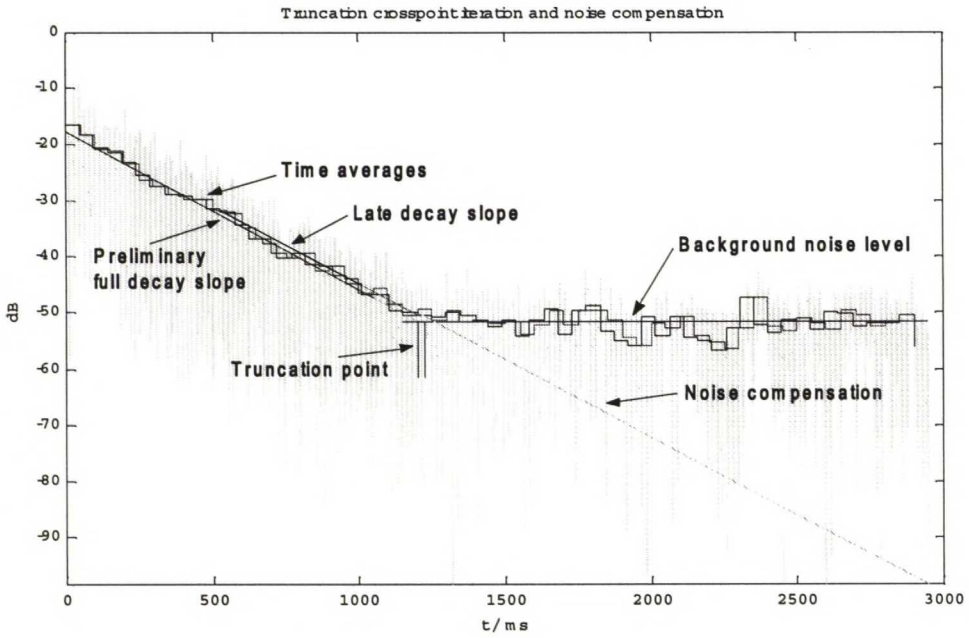


Figure 5.8. Operation of the Lundeby method.

6 Results and discussion

In this chapter, the results of the comparative measurements performed in Chapter 5 are presented and analyzed. This is followed by a discussion and evaluation of the IRMA measurement system in various views.

6.1 Comparison of results acquired with IRMA and MLSSA

First, the analysis methods applied on each system are discussed.

Results from the comparative systems are presented in linear and energy-time graphs of the wide-band impulse responses. These provide a basic but effective comparison of the equivalence between systems.

The results continue with graphs of room acoustic indices calculated for each response in third octave bands. These include the reverberation times T_{10}, T_{20} and T_{30}, clarity indices C_{50} and C_{80}, definition D_{50}, center time T_S, as well as the background noise levels.

In addition to results calculated on each system for its own responses, the acoustic indices also include a series of values, which were analyzed with IRMA using the MLSSA responses. This enables the differences in response processing and acoustical parameter analysis to be viewed.

A further comparison was made by calculating all results on the IRMA system with both forward and time-reversed filtering. This was to investigate how the filtering scheme may affect the results for each room acoustic criterion.

6.1.1 Analysis issues

IRMA

The IRMA system was used for analyzing both its own responses and the responses acquired with MLSSA. The latter was performed by reading the MLSSA impulse responses into Matlab, and accounting for the differing sample rates prior to analysis.

Analysis was performed in each case using both ordinary and time-reversed filtering schemes for all parameters.

All responses were first filtered, and then each band was individually post-processed to determine the background noise level, direct sound and noise floor crosspoint. Room acoustical indices were calculated from the processed data. The filter delays in the ordinary filtering scheme were accounted for in calculating the energy-time parameters.

MLSSA

The MLSSA system was used for analyzing its own responses. The noise compensation mode was turned on in the Schroeder plot view for both automatic and manual analysis, although it was not found to have an effect on the automatic analysis regardless of the claims in the online documentation. The automatic analysis was performed by using the Calculate Acoustics

command. In manual analysis, the parameters were determined by manually locating the reverberation level limits on the Schroeder integrated decay curve, and requesting the resulting slopes and correlations from the system.

6.1.2 Impulse responses and energy-time curves

The acquired impulse response and energy-time curves are presented below for both systems. Both the complete response and the first 100 ms of the response are shown for each pair.

Meeting room

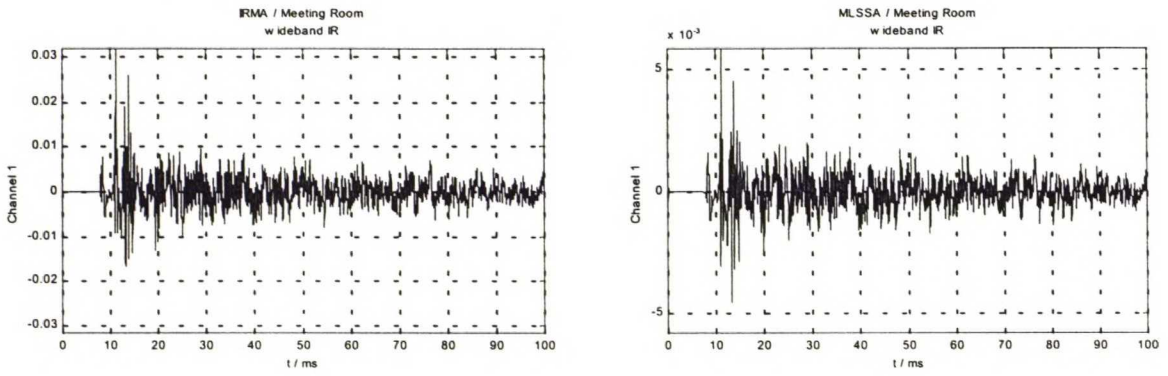


Figure 6.1. Direct sound and early reflections.

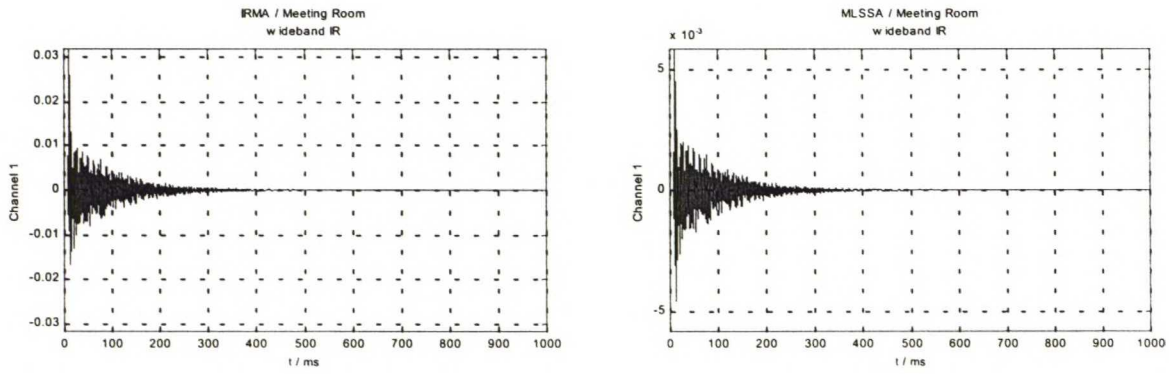


Figure 6.2. Complete responses.

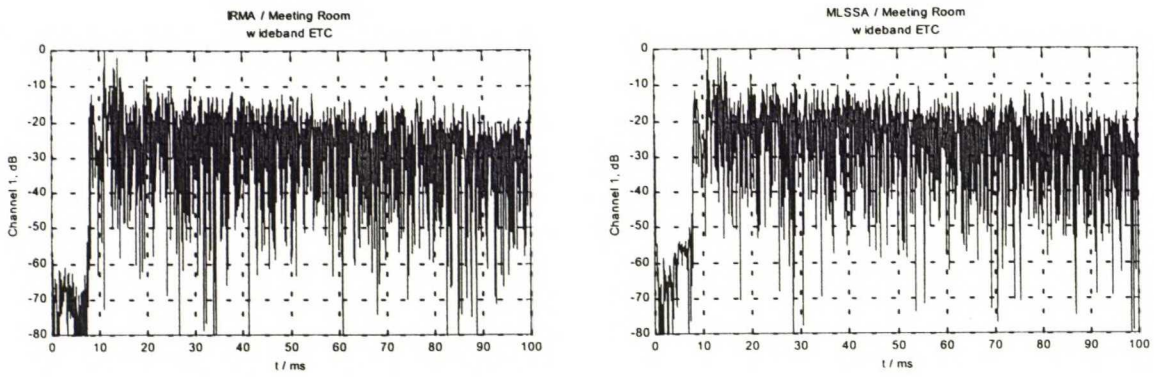


Figure 6.3. Energy-time plots of direct sound and early reflections.

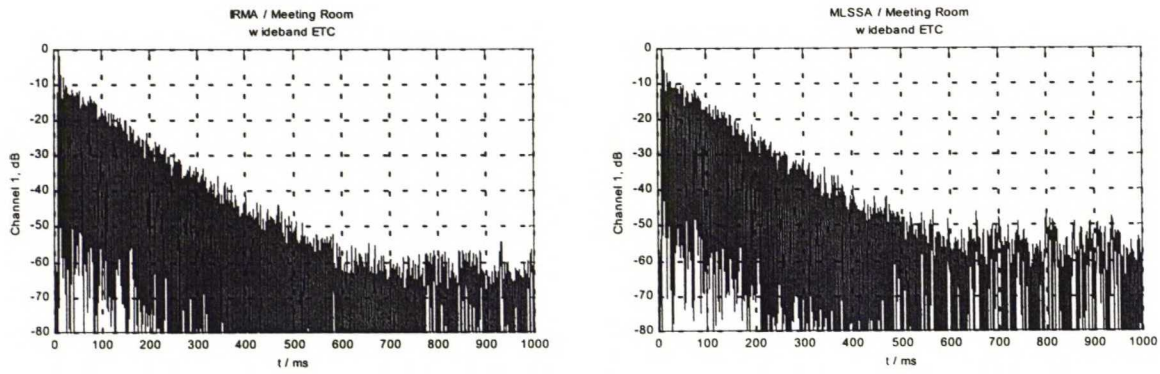


Figure 6.4. Complete energy-time plots.

The impulse responses acquired with the IRMA and MLSSA systems are found to be in good accordance with one another.

Small variations can be distinguished in the amplitudes of individual reflections. These may be due to the differences in the loopback impulse responses between systems. Similar differences between systems have been noted in literature [Lundeby *et al.* 1995]. Small random variations in the acoustic system are another possibility for these differences.

The background noise level is found to be better on the IRMA system. This is likely to be due to the 12 bit A/D resolution on the MLSSA system.

Concert hall

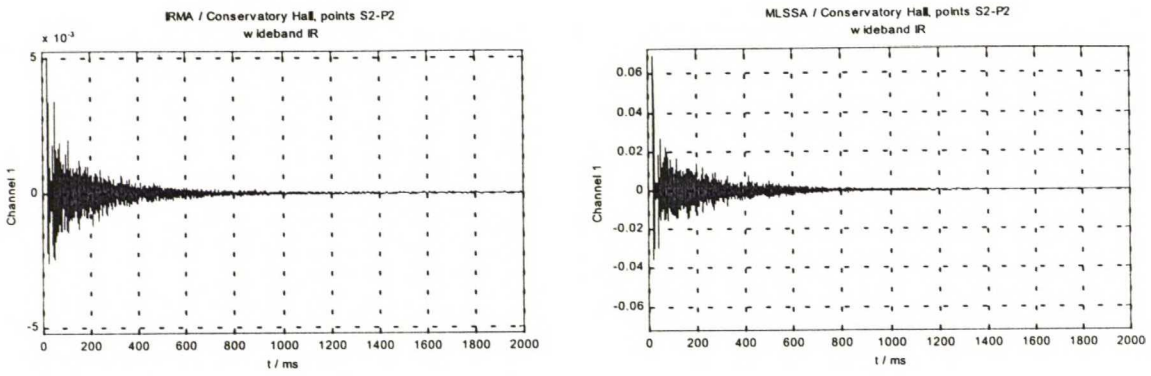


Figure 6.5. Complete responses.

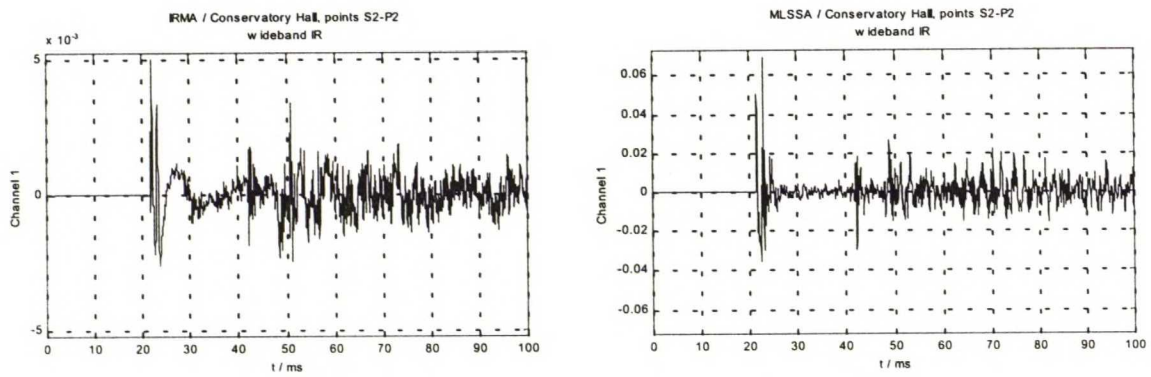


Figure 6.6. Direct sound and early reflections.

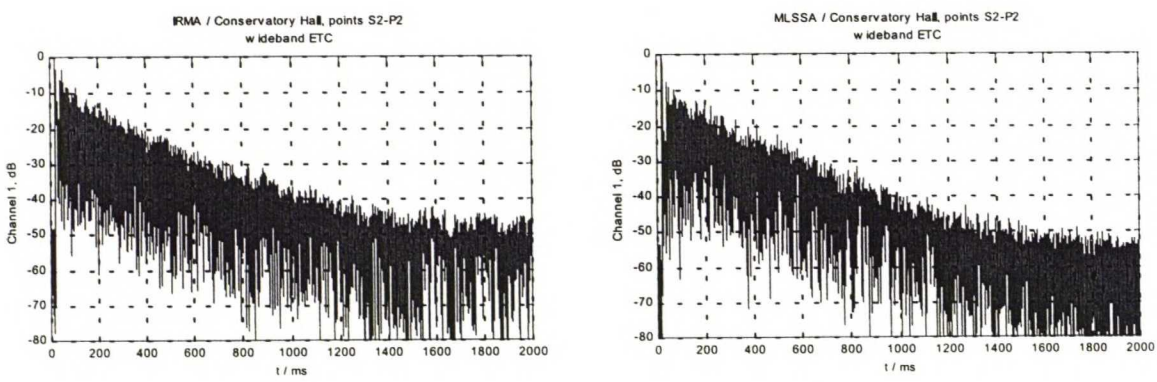


Figure 6.7. Complete energy-time plots.

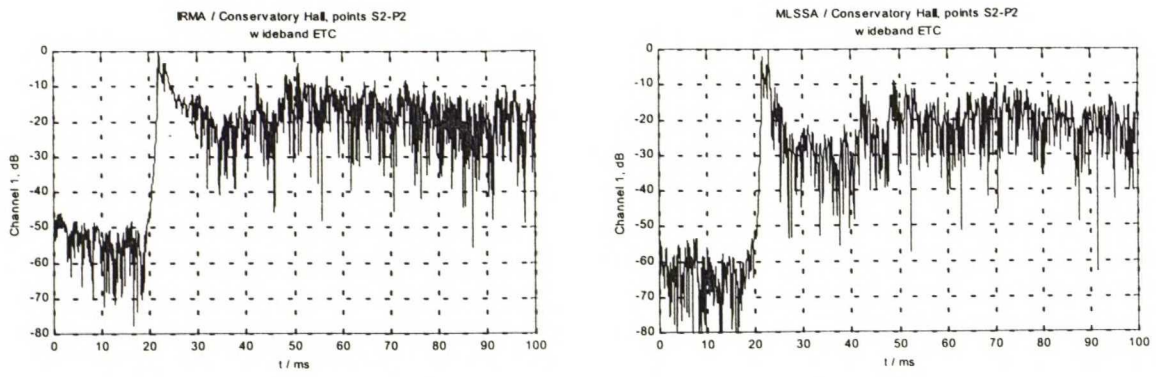


Figure 6.8. Energy-time plots of direct sound and early reflections.

The concert hall impulse responses show a considerable amount of differences in the smaller details. Both the linear response and EDT graphs show major differences after the direct sound. This behavior is likely to be caused by a variety of effects, as no simple reasoning can be given. The larger details of the responses are in accordance with one another.

Loudspeaker distortion forms one possibility of error for the IRMA response. The sluggish behavior after the direct sound, together with the increased background noise level are both factors, which could support this possibility [Vanderkooy 1994]. As the measurements were performed using an unfiltered MLS stimulus, it is very difficult to notice the presence of slight distortion in the stimulus signal.

6.1.3 Room acoustic indices

The room acoustic indices are presented in the following figures. Each figure depicts a single room acoustic criterion, with the various combinations of measurement, analysis and filtering listed in the legend.

Table 6.1. Combinations of response acquisition, analysis and filtering methods applied for the calculation of room acoustic indices.

acquisition system	analysis system	filtering method
IRMA	IRMA	ordinary time-reversed
MLSSA	MLSSA, automatic MLSSA, manual	(undocumented) (undocumented)
MLSSA	IRMA	ordinary time-reversed

Meeting room

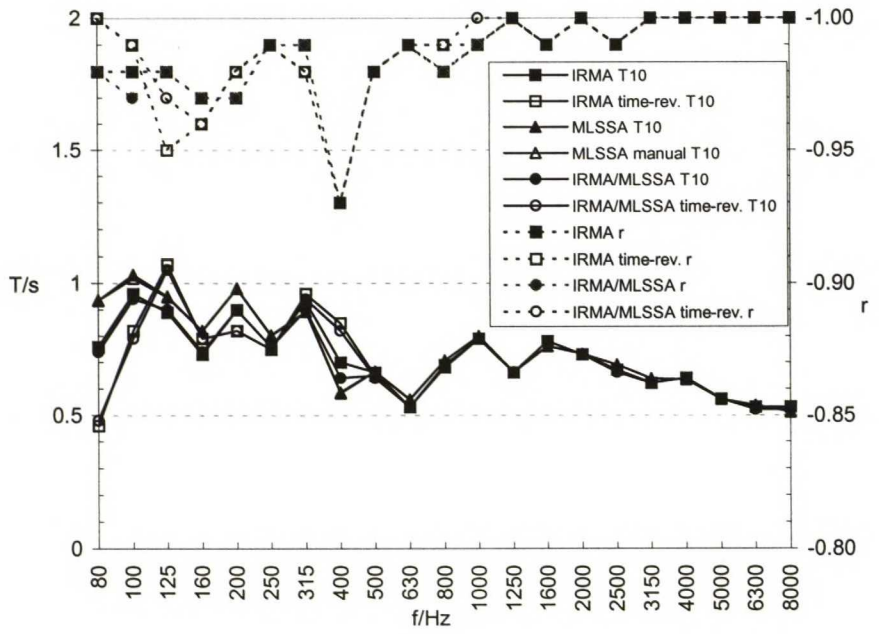


Figure 6.9. Early decay times and their correlations measured at the meeting room. Note that EDT correlation is not available on the MLSSA system.

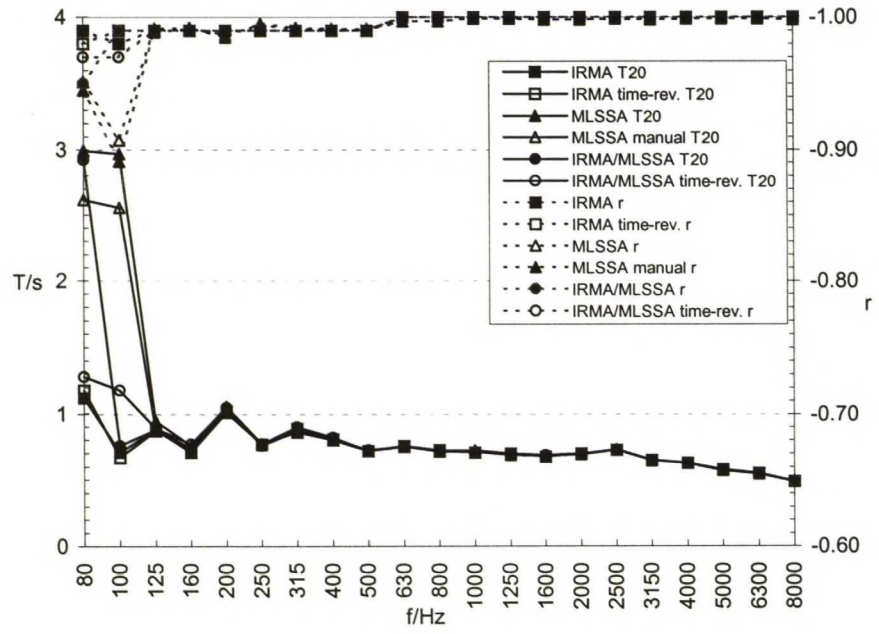


Figure 6.10. Reverberation times (T_{20}) and their correlations measured at the meeting room. Observe the expanded scaling to show the low frequency irregularities.

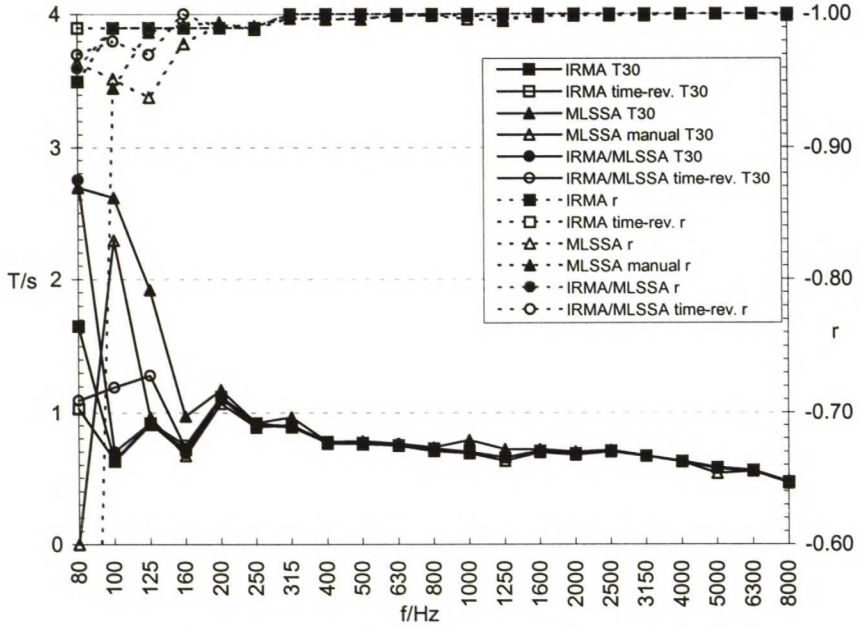


Figure 6.11. Reverberation times (T_{30}) and their correlations measured at the meeting room. Observe the expanded scaling to show the low frequency irregularities.

The reverberation times measured at the meeting room are generally in good accordance with one another, especially at frequency bands above 500 Hz.

At the lowest frequency bands, room modes cause distinct warble to the decay curves, resulting in naturally lower correlations for the decay approximations. The IRMA system seems generally more robust at low frequencies, yielding less erroneus results. This is most likely due to the advanced Lundeby algorithm that is used to determine the individual starting point and decay-noise kneepoint for each filtered response. However, MLSSA's impulse response is found to cause peaks at the lowest bands even when analyzed with the IRMA system.

As predicted, the EDT can be seen to fall at low frequencies, when time-reversed filtering is used.

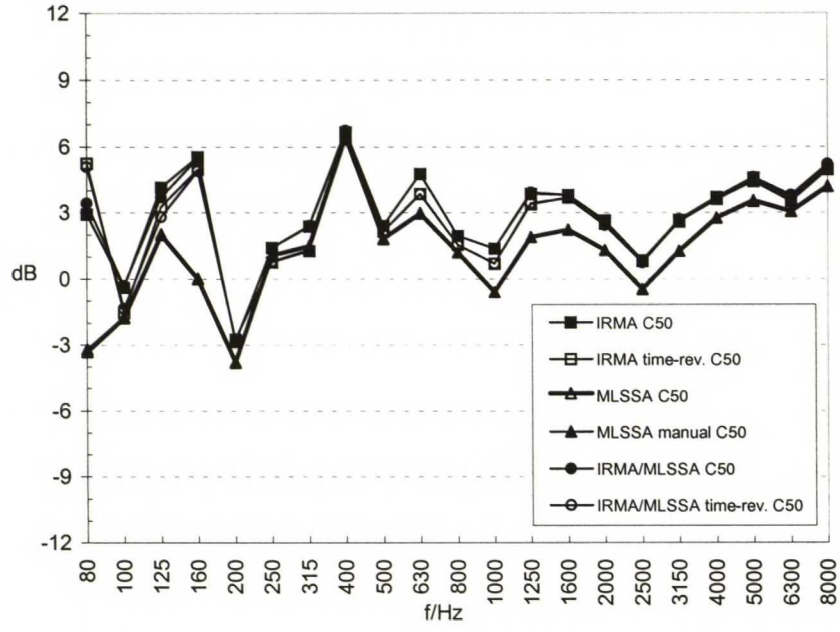


Figure 6.12. Clarity C_{50} measured at the meeting room.

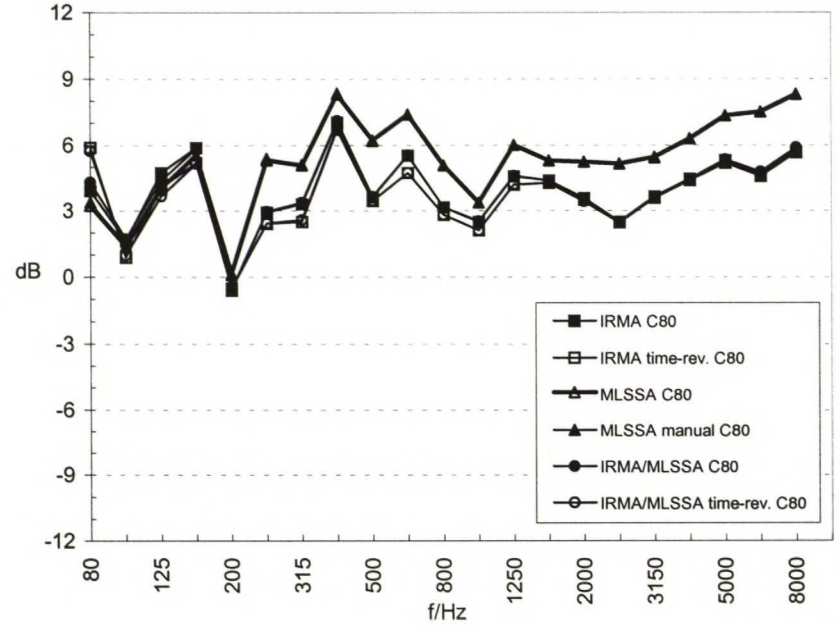


Figure 6.13. Clarity C_{80} measured at the meeting room.

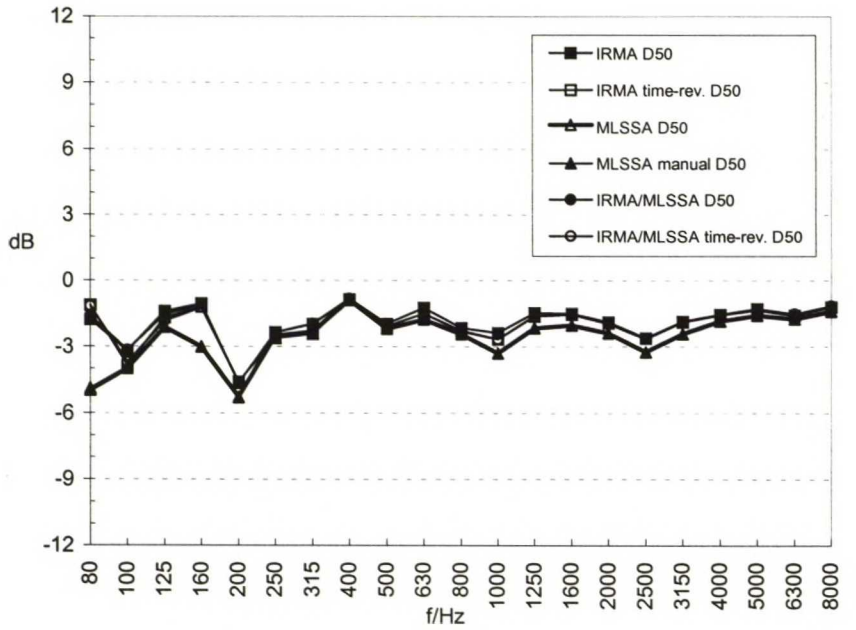


Figure 6.14. Definition D_{50} measured at the meeting room.

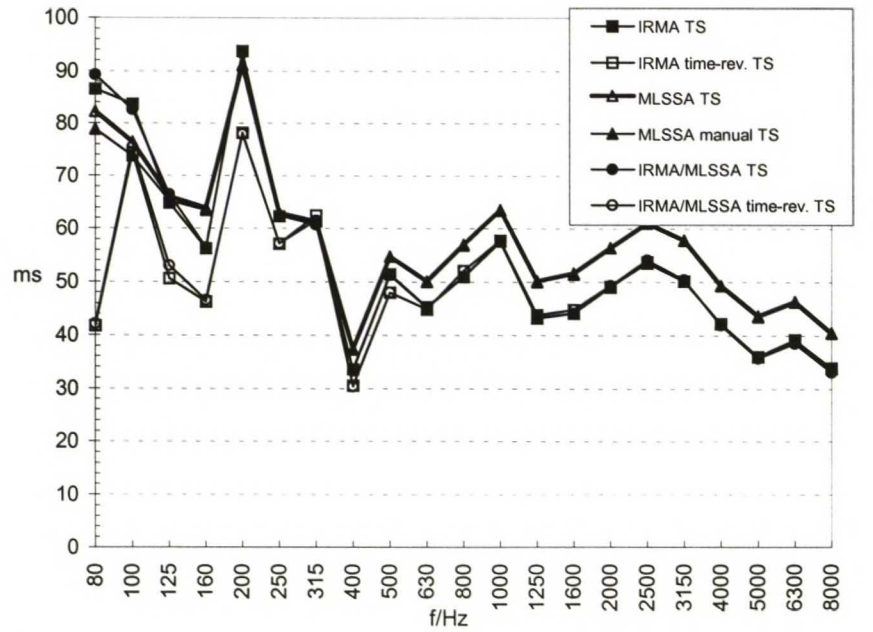


Figure 6.15. Center time T_S measured at the meeting room.

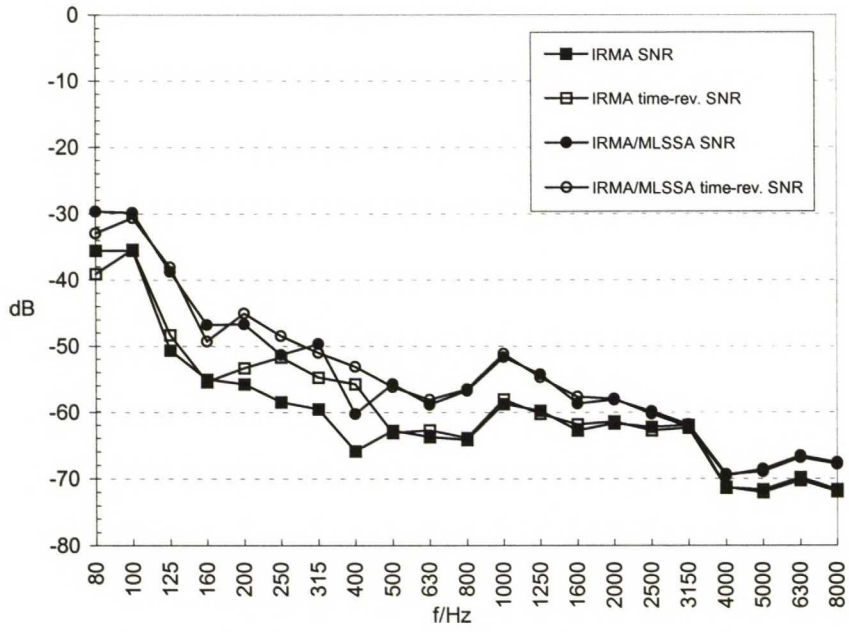


Figure 6.16. Impulse response background noise levels (using the direct sound peak for 0 dB reference) measured at the meeting room. This ratio is often denoted as the impulse response's SNR, although this is not a strictly correct use of the term.

The clarity and definition curves are found to have common form, but an offset of a few decibels is noted between systems. The same behavior is seen in the center time graphs, with the MLSSA results 7–8 ms above the IRMA results. This difference is very likely caused by the direct sound delay, which is of the same size. This would lead to the hypothesis, that the MLSSA system does not properly account for the direct sound delay in the analysis, thus causing a systematic error in the results.

The background noise levels have been analyzed for both IRMA and MLSSA responses using the IRMA system for analysis. Again the results suggest, that the IRMA system has a wider dynamic range in measurements.

Concert hall

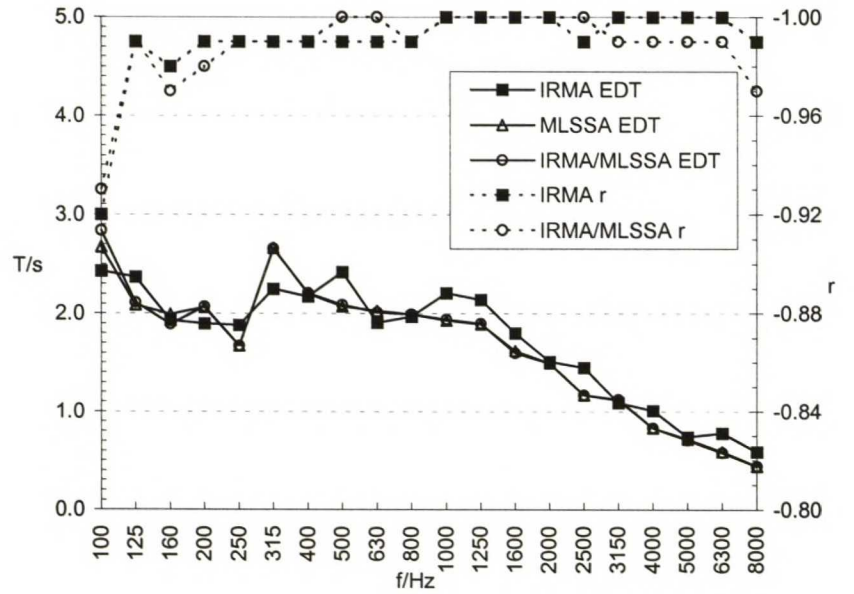


Figure 6.17. Early decay times and their correlations measured at the concert hall. Note that EDT correlation is not available on the MLSSA system.

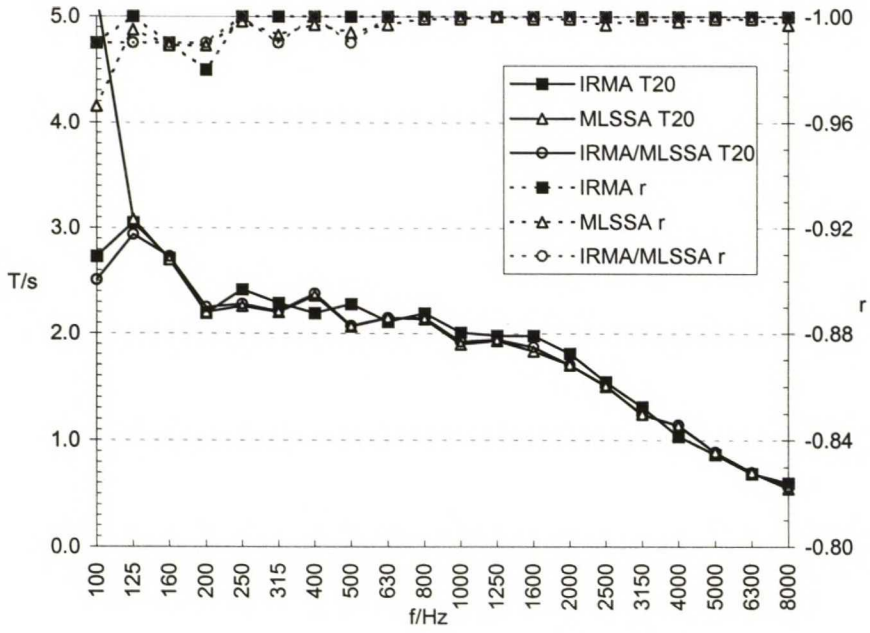


Figure 6.18. Reverberation times (T_{20}) and their correlations measured at the concert hall.

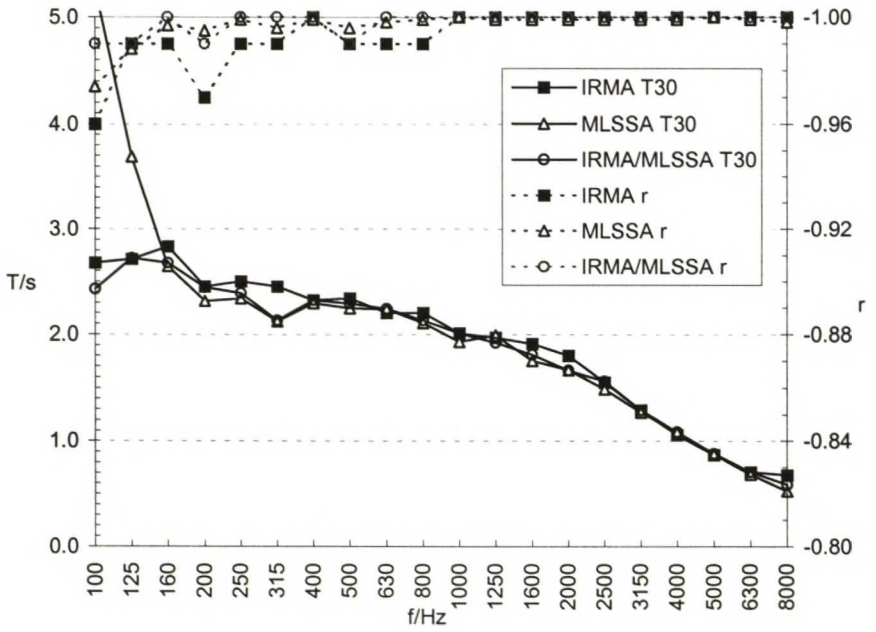


Figure 6.19. Reverberation times (T_{30}) and their correlations measured at the concert hall.

The reverberation times T_{20} and T_{30} measured at the concert hall are in good accordance with one another, except for the 100 and 125 Hz bands, where the MLSSA system's analysis fails. The EDT results show a larger amount of variance, although the results between systems are well within the repeatability limits of a typical reverberation measurement.

With reverberation times several times larger than at the meeting room, the differences caused by time-reversed filtering are found negligible.

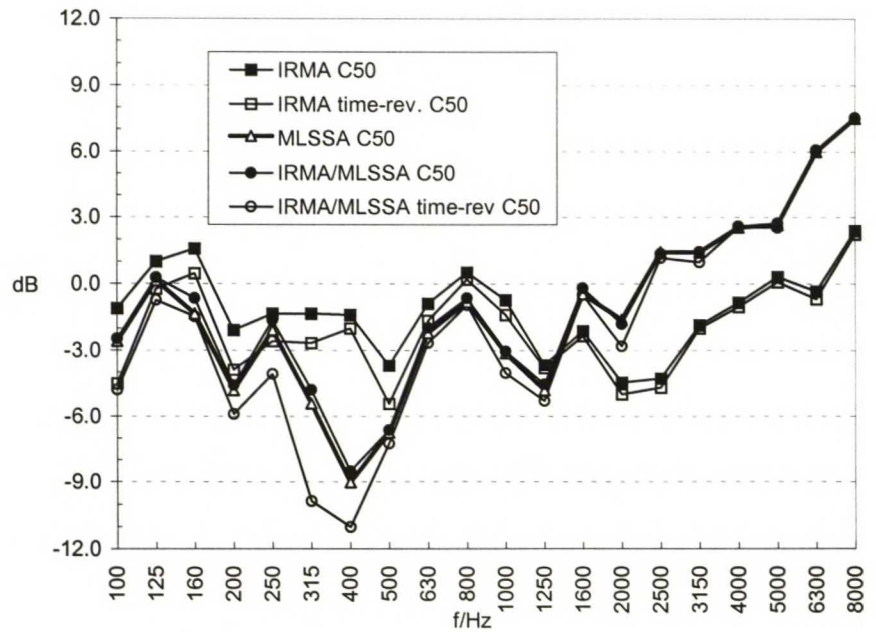


Figure 6.20. Clarity C_{50} measured at the concert hall.

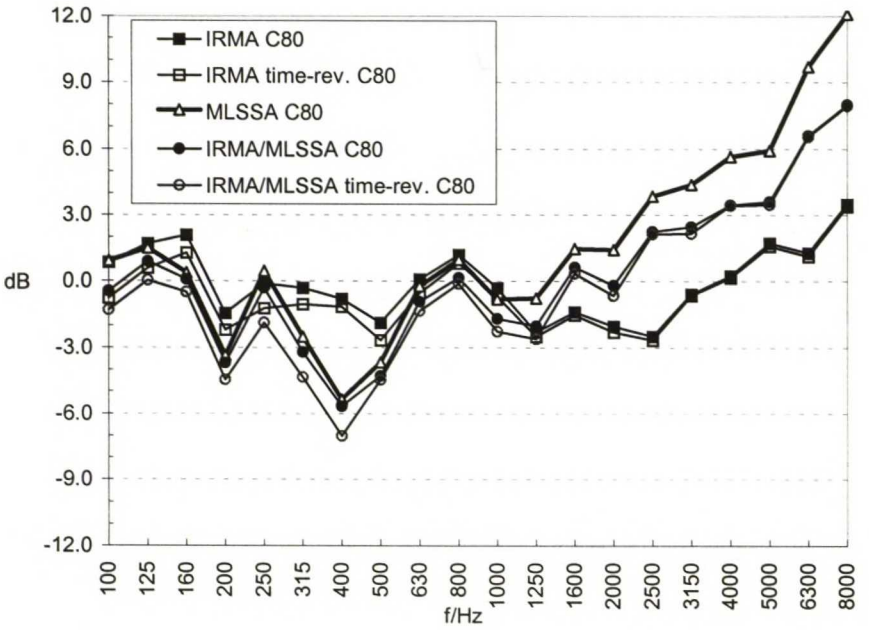


Figure 6.21. Clarity C_{80} measured at the concert hall.

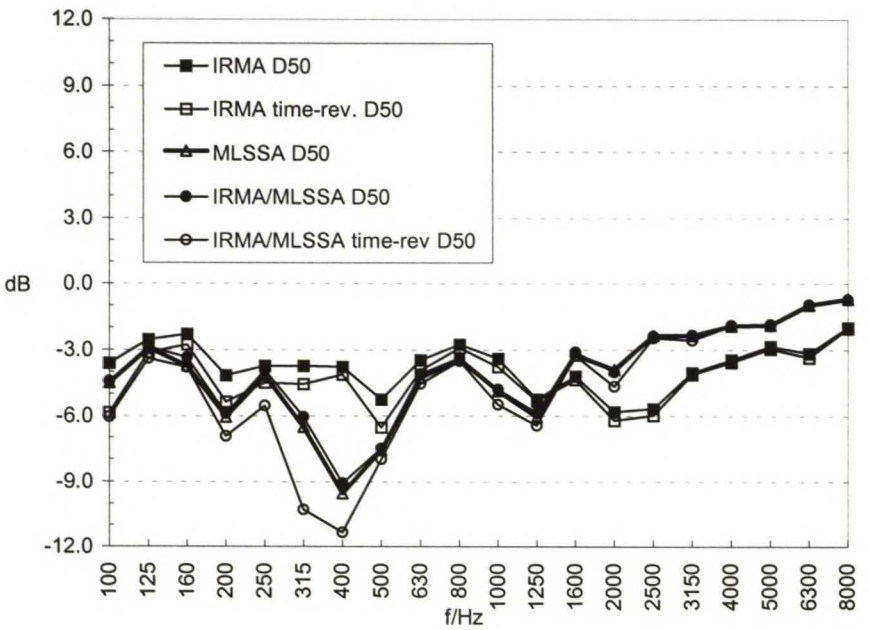


Figure 6.22. Definition D_{50} measured at the concert hall.

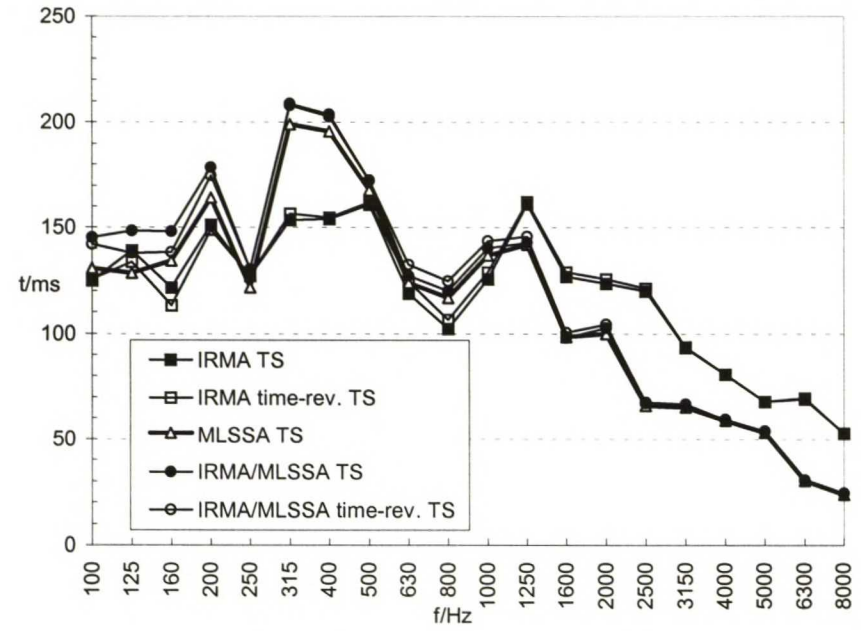


Figure 6.23. Center time Ts measured at the concert hall.

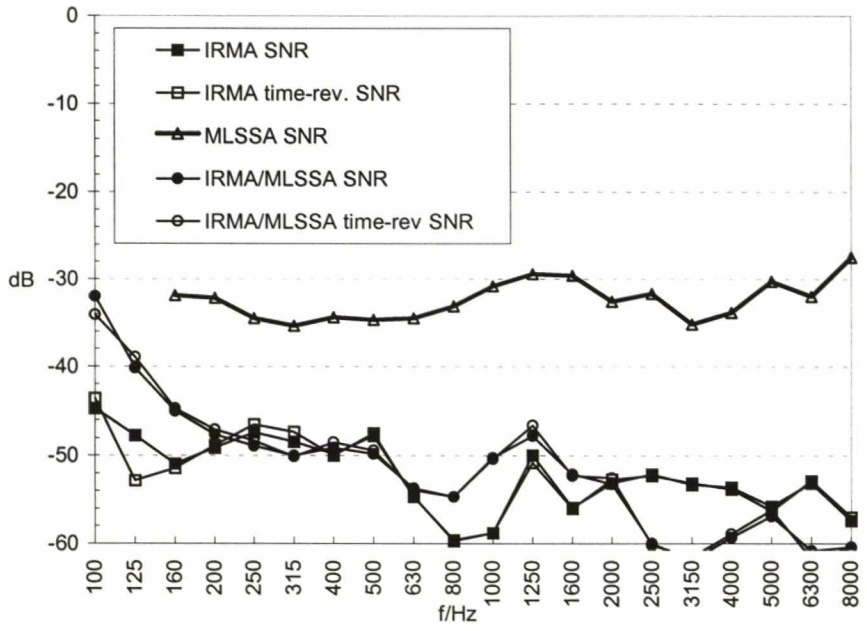


Figure 6.24. Impulse response background noise levels (using the direct sound peak for 0 dB reference) measured at the meeting room. This ratio is often denoted as the impulse response's SNR, although this is not a strictly correct use of the term.

When compared to the wide-band direct sound delay of about 22 ms, the center time results repeat the notion of inferior direct sound detection in the MLSSA system at bands above 1 kHz.

Differences between the two impulse responses are clearly visible in the clarity and definition results, where the IRMA response results differ from the MLSSA response results calculated with either system. The C_{80} graph exhibits an exceedingly large variance between the different results, with spreading between all three acquisition–analysis system combinations.

The adverse effects of time-reversed filtering for energy-time analysis are clearly seen at the lower mid-frequency area, with differences of 1–2 dB between the filtering methods.

The background noise levels portray again a larger dynamic range for the IRMA system at the low and mid-frequency bands. The high-frequency behavior may be due to loudspeaker distortion having occured in the IRMA response measurement. The MLSSA SNR value is included only for curiosity, as the method of calculation applied by the system is undocumented. Clearly it is not comparable to the other results.

6.1.4 Sources of error

In both measurements, the absolute stimulus amplitudes of the systems were not calibrated with each other. This forms a potential source of error for the inherent system noise floor determination. As the measured noise floors are mostly above the systems' electrical background noise floors, it is likely, that the noise present in the measured responses may be for some part accounted for as acoustical background noise.

However, slight distortion in loudspeaker reproduction may well also be the cause for background noise in the –50 to –60 dB range. Loudspeaker distortion could also be a possible cause for the differences accounted for in the responses measured at the concert hall.

External low frequency disturbances, such as randomly varying noise emitted by the ventilation system or outside traffic may be another cause for error. In smaller amounts, this commonly causes variance in the background noise floor level and shape when using the MLS method [Nielsen 1996].

A limiting factor for the concert hall measurement was also the fact that the MLSSA response bandwidth was limited to 10 kHz, whereas the IRMA response has a bandwidth of 20 kHz. This is due to the MLSSA system, which is not capable of producing MLS stimuli of orders above 16. Thus the sample rate has to be reduced for acquiring long responses with MLSSA, effectively also reducing the measurement bandwidth. This can account for some of the differences observed between the response curves. Naturally, this difference does not affect the results calculated from band-filtered data.

6.2 Discussion on the IRMA system

The advantages and drawbacks of the IRMA system are discussed, with regard to hardware, software and method-related issues.

6.2.1 Hardware issues

Microphones

High quality measurement microphones commonly exhibit highly linear frequency and phase response characteristics, along with low noise and distortion figures. However, there are several properties of measurement microphones that should be taken carefully into account in room acoustical measurements.

With the exception of small omnidirectional capsules, microphone directivity is not always ideal at all frequencies. All large microphone capsules exhibit inherent phase limitations at high frequencies. Directional microphones with figure-of-eight or cardioid patterns are an important and often neglected case, where the directional patterns commonly exhibit major variations at low-frequency bands.

Microphone sensitivity is a factor where notable variations are often present even between two individual microphones of the same make and model. Sensitivity does not have an effect on simple response measurements, but it should be noted whenever absolute or relative levels are required, such as in IACC, LEF and sound strength measurements.

The accurate measurement of LEF values requires that the omnidirectional and figure-of-eight microphones be calibrated for equal sensitivity. The figure-of-eight microphone should be calibrated in the most sensitive directions of the capsule. This must be done in a free field using the comparison method, as there are no designated calibrator units suitable for calibrating figure-of-eight capsules. Even with calibrated microphones, a definite source of error are the variations in sensitivity in all the other directions between different figure-of-eight capsules and types. This makes it difficult to accurately compare LEF measurements made with different apparatus.

Loudspeakers

Loudspeaker units are commonly the main cause for distortion in acoustical measurement systems. High output levels cause a multitude of nonlinear effects in common dynamic loudspeaker units, such as unlinear cone motion due to mass and resonance effects, magnetic effects due to the mechanical offset, heating and saturation of the voice coil, etc. Many modern loudspeaker units[1] also include active protection circuits, which compress and limit the output signal at high levels.

[1] For example, the Genelec 1029A and 1030A active loudspeaker models used in the testing phase.

These nonlinearities may generally be reduced by using sufficiently powerful equipment, and by lowering the stimulus signal levels with either band-limited or low-pass filtered stimuli, or by employing noise tolerant measurement methods.

Analog signal conditioning

In general, the nonlinearities produced by microphone preamplifier and power amplifier units under proper operating conditions are of such a low magnitude, that they are unlikely to stand out in the measurement chain.

Amplifier background noise may become a relevant factor at low levels with high signal gains. Exceedingly high signal levels may introduce clipping, which results in drastic distortion values. Under mindful operation these effects are seldom a problem.

A/D and D/A converter impulse responses

When a digital unit impulse is converted to analog and then back to digital form, the bandlimiting anti-alias filtering cause ringing in the form of a sinc (*i.e.* $\sin(x)/x$) pulse in the system response. Depending on the architecture of the converters and anti-alias filters, the loopback response may be symmetric or asymmetric in time, and vary in length and shape. This response is inherent to the system, and affects all the impulse responses acquired with the system.

Acquired impulse responses can be compensated by deconvolving them with the system's electrical loopback response. For room acoustic purposes, the time scale of interest is usually much larger than the electrical system response length, so compensation is not always necessary. However, measurements of electrical impulse responses as well as anechoic acoustic impulse response are affected.

The Korg audio hardware was noted to introduce a symmetric sinc pulse with a total length of a few milliseconds, when the system analog output was looped directly back into an analog input. For room acoustic measurements and analysis, this time scale does not cause major inaccuracies. Applications employing short impulse responses, such as anechoic speaker response measurements may be affected, though. The ringing effect may be surpressed either by using loopback compensation by deconvolution, or by lowpass filtering the response to a smaller bandwidth.

For comparison, the MLSSA system uses a 12-bit successive approximation A/D converter, which causes asymmetric ringing only after the arrival of the impulse. Together with programmable antialias filters and support for high sample rates the system exhibits neglectably short ringing, the length of which is in the order of 0.1 ms.

Currently, the MLSSA system is better suited for measurements of short responses. However, the S/PDIF and T-DIF connections in the IRMA system hardware make it possible to use any external A/D converter unit with a more suitable ringing performance. Unfortunately, the manufacturers of studio sound equipment seldom document the type of converters employed in their converter units. Thus the ringing properties of a converter unit cannot usually be judged without testing the device.

Background noise floor levels

Factors affecting the background noise floors in loopback responses include A/D and D/A converter bit resolutions, the amount of thermal noise and disturbances in the analog signal chain. In practice, converters with more bits do not always yield a better noise figure in the measurements. The MLSSA system, for example, [Rife 1996] utilizes a 12 bit A/D converter, but with an automatic gain control preceeding the converter input, the system can produce loopback SNR values exceeding 70 dB.

Computer sound cards

Another background noise factor common to many ordinary PC sound cards is irregular time invariance, which can result in adverse SNR ratios for periodic impulse responses, such as by employing the MLS method.

There are various possible reasons for this behavior, such as bad synchronization between the input and output signals, flow errors in data buffering. Curiously, many modern sound cards also attempt to enhance the output signal by various methods, such as virtual surround effects etc. Commonly the processing involved in these features is undocumented, and it may even be impossible to bypass these effects from the signal chain. It is clear that this kind of hardware should be avoided for measurement purposes.

In addition, the wealth of digital interference sources inside a computer makes the operating environment of an internal computer sound card problematic with analog audio signals. Although high quality sound cards are generally well protected against EMI, sharp narrow-frequency spikes of interference may still exist especially in the input signals. From the EMI standpoint, a better alternative is to keep the analog audio signals outside the computer chassis by use of external converter units.

Synchronization issues

Time synchronization of the input and output data is of great practical importance in impulse response measurements. In order to gain time-accurate responses, there should be no mutual delays between the signal inputs and outputs of the system. In practice, interchannel synchronization is mainly affected by two factors: sample clock timing and data buffering.

In computer sound cards, sample clock timing is not an issue, as the on-board converters share a common clock source. With external converter units, it is important that all the separate digital audio components are synchronized to the same clock source. Separate clock signal lines from one unit to another help to avoid jitter problems.

Data buffers are another source of delay, but they are necessary to ensure the real-time transfer of signal data without dropouts. Data buffering usually takes place both in the sound hardware and the driver software. Under the Windows architecture, common sound cards are found to exhibit loopback buffer delays in the order of a few hundred samples. If there is no need to trigger the measurement from an external source with accurate timing, the static time lag caused by the buffering does not usually cause any practical problems.

A difficult source of error results from interchannel delays that exhibit random deviations. A static delay is easily compensated, but random deviations in the delay make it impossible to accurately predetermine the starting point of a measured impulse response. This was found to be a common problem with low-end consumer sound cards.

The clock sync connections on the Korg 1212 I/O sound card enable the whole measurement system to be connected to an external clock source. This would be applicable for time-interpolated MLS measurements in the ultrasound range [Mommertz, Bayer 1995].

Level control

The lack of computer control over the analog input gain settings can be a major disadvantage, if highly automated measurements are required. The measurement application can monitor the digital input levels and inform the user of the peak levels and possible clipping in the inputs, but it has to rely on the user to adjust the input gains accordingly during a measurement session. With a computer sound card, the analog input gain levels may usually be controlled by software.

Although the Korg 1212 I/O card has standard Windows sound drivers, the mixer settings have not been implemented according to the Windows sound API, but use a custom application instead. The analog input levels on this card cannot thus be directly controlled by the custom measurement software. For most two-channel sound cards, this approach could readily be employed.

Level calibration

Level calibration issues pose another problem with external AD converter units. Accurately reproducible input gain settings require level controls with discrete stepped values. Unfortunately, all the suitable AD converters available at the time of conducting this work employed continuous range potentiometers, as the units are inherently designed for the studio market. Thus a compromise had to be taken, and the basic Korg 880 A/D unit was chosen.

There are several solutions to the calibration issue. All the input gain controls can of course be set to full level, and use the microphone amplifier for adjusting the gain. If this results in an excessive input gain, it could be fixed to another value by replacing the converter unit's gain potentiometers with discrete precision resistors.

Input level calibration

Input level calibration implies a need for controlling the gain of the measurement system's analog inputs in discrete steps, so that the same input gain could be used for different channels and measurement sessions. This need encompasses the whole analog signal chain, comprising of the measurement microphones, the microphone preamplifiers and the AD converters. Effectively it would be optimal, if the analog gain levels could be programmed and stored automatically by the measurement computer, making the tasks of performing and analysing measurements both faster and less prone to user errors.

6.2.2 Software issues

The Matlab environment

The Matlab environment is well suited for the analysis and visualization of large amounts of data. Development and coding under Matlab is rapid and easy when compared to standard lower level programming languages, such as C(++). This is largely due to the abundant functions and matrix operations contained in Matlab.

However, the Matlab environment has its drawbacks, when excessively large amounts of data have to be handled. There is an excessive resource overhead both in the practical storage space required for numerical data regardless of its accuracy, as well as the radical decrease in computational speed for certain types of repetitive functions, which cannot be reduced to basic matrix operations. Also, the inability to pass arguments by reference to functions was found problematic in more complex applications, as the memory overhead taken up by redundant copies of arguments in nested functions can grow very large.

The workarounds for these issues are either the extensive use of C coded Matlab mex files, or a large amount of effort in coding to reduce allocation overhead by other means. Both methods were used in certain parts of the IRMA software.

The MLS generator and time-domain deconvolver functions were coded as mex files, as their Matlab counterparts would have been unfeasibly slow. Functions for reading and writing wave files to disk were also replaced with mex files, with a resulting speed increase of several decades.

The Matlab base workspace and the global variable workspace were used between some functions for storing large amounts of data, with only variable names passed as arguments between functions. This is a form of passing arguments by reference, but the use of global or base workspace variables for passing data between functions cannot be considered a very solid or safe coding practice.

The use of the Matlab environment for real-time data acquisition is another question of debate. The Data Acquisition toolbox was tested for suitability as a replacement for the recplay application, but the system was found to yield variable latency times between the input and output channels.

In order to ensure errorless measurements, the realtime data acquisition must take place independent of the Matlab environment. However, it is quite possible, that a mex file version of the recplay application could be implemented virtually inside Matlab.

The IRMA software

The IRMA software has been found well capable of practical multichannel measurement use. Response measurements, plotting and storage can be rapidly done in the field.

The data analysis part has also been found to yield accurate and correct results. However, it is more involved than the measurement part, as the forms of compensation, filtering and processing required prior to analysis vary in each

case. Typically the analysis of room acoustic parameters requires separate processing for each filter band of each channel in the response. In effect, the amount of data and storage involved quickly multiplies to cumbersome proportions.

The implemented basic data visualization includes plotting functions for time, energy-time and Schroeder integrated responses, as well as frequency response.

For rapid evaluation of acoustical parameters from measured responses could be done with a separate graphical analysis interface, with seamless integration to the filtering, processing and acoustical analysis functions. This would enable the fast evaluation and comparison of decays and acoustical parameters at different frequency bands, preferably for a number of channels in any acquired response.

Development of the IRMA software is likely to continue after this work. The key points will likely be further development of data analysis, as well as enhancements to the current user interface. Also, the possibility to select between parallel and serial operation of using channels would give the user the choice between processing time and memory allocation for the analysis of multichannel and multiband responses.

7 Conclusions

The scope of this work was to design and build a multichannel system for the measurement and analysis of acoustical impulse responses and to accomplish an insight on the subject of current room acoustical measurements and the analysis methods involved.

Acoustical impulse responses provide the basis for calculating standard room acoustical parameters, such as reverberation and energy-time relations, as well as directional indices. The acoustical properties of spaces may be studied in more detail by using multichannel impulse response techniques, which provide accurate directional or locational information on the sound field.

The MLS method is well suited for the determination of impulse responses in linear time-invariant systems. It was used as a basis for the measurement system.

The IRMA measurement system was put together as a part of this work. The system consists of a Matlab-based application for response measurement and analysis, and a portable measurement computer with multichannel audio hardware. A functional layout of the system is depicted in figure 6.1.

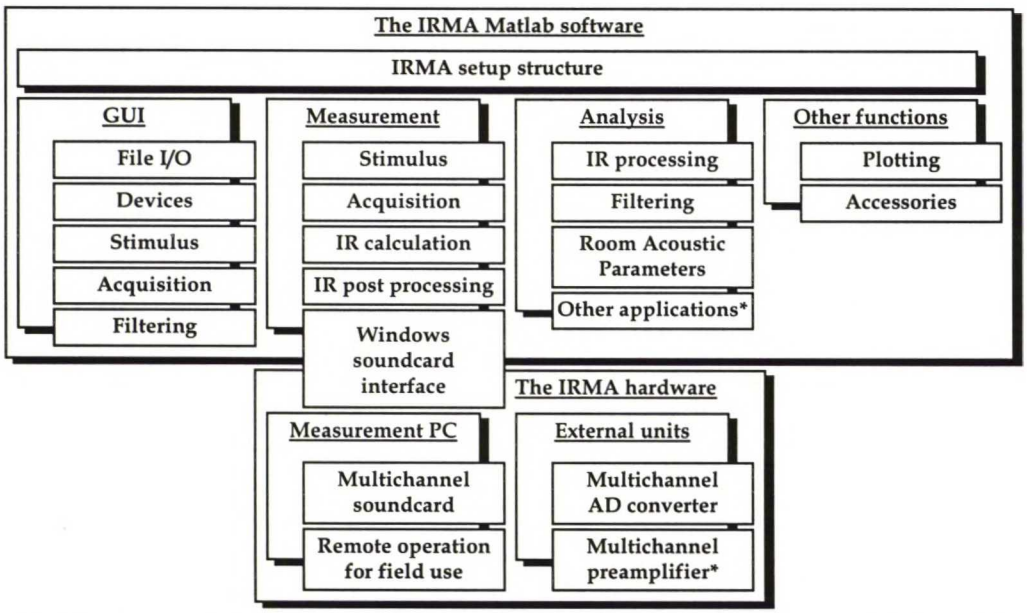


Figure 6.1. A complete functional layout of the IRMA measurement system.

Comparative impulse response measurements and room acoustical analysis were carried out using the IRMA system and the commercial MLSSA system. The measurements were performed at an office meeting room and at a concert hall. Comparison of results shows that the systems agree well with one another, with minor differences between the responses and room acoustical parameters analyzed. The IRMA system's analysis methods are found to be robust when compared to the MLSSA system.

Although the impulse responses acquired with different measurement systems may be very similar, there can be large variance in the methods applied for filtering and processing responses prior to the analysis of room acoustical parameters.

Bandpass filtering of responses into octave or third octave bands causes a time-spreading of signal energy due to filter delay and decay times. Time-reversed filtering may be used to minimize this effect for the accurate determination of short reverberation times. However, EDT and energy-time ratios should be analyzed using ordinary filtering and compensation for filter delay, in order to avoid initial time-smearing of the response.

Measured impulse responses must be processed to remove unidealities prior to calculating room acoustical parameters. These parameters are based on an ideal exponential decay model, so the background noise and initial delay present in real responses must be removed prior to analysis. Various methods have been studied for this purpose. The algorithm described by [Lundeby *et al.* 1995] has been implemented in the IRMA system.

In general, MLS-based measurement systems produce accurate results with small variance between one another. However, the MLS method is quite sensitive to nonlinearities or time invariance present in the measurement signal chain. Factors such as loudspeaker distortion or poor digital synchronization may have very adverse effects on the results.

MLS-based measurements should be used with care and a trained eye for potential misbehavior. This ensures fast acquisition of optimal and accurate results with a large dynamic range.

In the near future, the IRMA system will be used for multichannel measurements at a number of Finnish concert halls. In addition to standard room acoustical measurements, a binaural head and a three-dimensional probe with 12 microphones are used to gather a representative amount of response data at each hall. It is very likely that the IRMA system will also be employed in analyzing the acquired data, with a focus on novel methods for studying the directional information of the sound field at the concert halls.

References

Ando Y. 1985. *Concert Hall Acoustics*. Springer-Verlag Berlin. pp. 102–114, 134–138.

Aurora 2000. The Aurora Homepage.
URL: <http://www.ramsete.com/aurora/home.htm>. 20.11.2000.

Barron M., Marshall A. H. 1981. Spatial Impression Due to Early Lateral Reflections in Concert Halls: The Derivation of a Physical Measure. *J. Sound Vib.* (1981) **77** (2). pp. 211–232.

Barron M. 1984. Impulse response teksting techniques for auditoria. *Appl. Acoustics*, Vol. **17**, 1984. p. 165. (referenced in [ISO 3382, 1997])

Barron M. 1995. Interpretation of Early Decay Times in Concert Auditoria. *Acustica,* Vol. **81** (1995). pp. 320–331.

Bendat J. S., Piersol A. G. 1980. *Engineering applications of correlation and spectral analysis*. Wiley. 302 p.

Beranek, L. 1988. *Acoustic Measurements*. Acoust. Soc. Am.

Beranek, L. 1996. *Concert and Opera Halls and How They Sound*. Acoust. Soc. Am.

Borish J., Angell J. B. 1983. An Efficient Algorithm for Measuring the Impulse Response Using Pseudorandom Noise. *J. Audio Eng. Soc.*, Vol. **31**, No. 7, 1983 July/August. pp. 478–487.

Borish J. 1985. Self-Contained Crosscorrelation Program for Maximum-Length Sequences. *J. Audio Eng. Soc.*, Vol. **33**, No. 11, 1985 November. pp. 888–891.

Bradley J. S. 1986. Auditorium acoustic measures from pistol shots. *J. Acoust. Soc. Am.* **80**(1), 1986 July. pp. 199–205.

Bradley J. S., Halliwell R. E. 1992. Sources of Error in Auditorium Acoustics Measurements. ICA 14th Preprint F3-1.

Brüel&Kjær 1992. *Technical Documentation. Two-microphone impedance measurement tube type 4206*. Nærum, Denmark. 87 p.

Carlson A. B., 1986. *Communication Systems*. Third Edition. McGraw-Hill. 686 p.

Chu W. T. 1978. Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method. *J. Acoust. Soc. Am.* **63**(5), May 1978. pp. 1444–1450.

Couvreur C. 1997. *The Octave Toolkit. A series of Matlab m-files generates A-weigthing, C-weighting, octave and one-third-octave digital filters.* URL: <http://ftp.mathworks.com/pub/contrib/v5/signal/octave/>. 20.10.2000.

Dunn C., Hawksford M. O. 1993. Distortion Immunity of MLS-Derived Impulse Response Measurements. *J. Audio Eng. Soc.*, Vol. **41**, No. 5, 1993 May. pp. 314–335.

Dunn C., Rife D. D. 1994. Comments and replies on "Distortion immunity of mls-derived impulse response measurements". *J. Audio Eng. Soc.*, Vol. **42**, No. 6, 1994 June. pp. 491–497.

Fahy F. J. 1989. *Sound Intensity*. Elsevier Science Publishers Ltd. 274 p.

Faiget L., Ruiz R., Legros C. 1996. Estimation of Impulse Response Length to Compute Room Acoustical Criteria. *Acustica* Vol. **82** (1996) Suppl. 1. p. S148.

Farina A., Tronchin L. 1998. 3D impulse response measurements on S. Maria del Fiore Church, Florence, Italy. *Proc. ICA 16th*. pp. 2465–2466.

Fausti P., Farina A., Pompoli R. 1998. Measurements in opera houses: comparison between different techniques and equipment. *Proc. ICA 16th*. pp. 679–680.

Gade A. C., Rindel J. H. 1984. *Akustik I Danske koncertsale ("The acoustics of Danish concert halls")*. The Acoustics Laboratory, Technical University of Denmark. Publ. No. 22, 1984.

Gade A. C. 1982. *Subjective room acoustic experiences with musicians.* The Acoustics Laboratory, Technical University of Denmark. Report No. 32, 1982.

Gade A. C. 1989. *Acoustical survey of eleven European concert halls—a basis for discussion of halls in Denmark.* The Acoustics Laboratory, Technical University of Denmark. Report No. 44, 1989.

Gade A. C. 1992. Practical aspects of room acoustic measurements on orchestra platforms. *Proc. ICA 14th.*, preprint F3-5. Beijing 1992.

Golay M. J. E. 1961. Complementary series. *IRE Trans. Info. Theory*, no. 7, pp. 82–87.

Gumas C. C. 1997. A century old, the fast Hadamard transform proves useful in digital communications. *Personal Engineering*, November 1997. pp. 57–63.

Halmrast T., Gade A., Winsvold B. 1998. Simultaneous measurements of room-acoustic parameters using different measurement equipment. *Proc. ICA 16th*. pp. 2467–2468.

Heyser R. C. 1988. *An anthology of the works of Richard C. Heyser on measurement, analysis and perception.* AES Inc., New York 1988. 279 p.

Hidaka T., Beranek L., Toshiyuki O. 1995. Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls. *J. Acoust. Soc. Am.* **98**(2), Pt. 1, 1995 August. pp. 988–1007.

Hirata Y. 1982. *J. Sound Vib.*, vol. **82** pp. 593–595.

IEC-1260:1995. *Electroacoustics – Octave-band and fractional-octave-band filters.* First Edition 1995-07. International Electrotechnical Commission. Geneve. 53 p.

ISO 3382:1997. *Acoustics—Measurement of the reverberation time of rooms with reference to other acoustical parameters.* International Standards Organization. Geneve. 21 p.

Jacobsen F., Rindel J. H. 1987. Time Reversed Decay Measurements. *J. Sound. Vib.*, Vol. **117** (1987) , No. 1. pp. 187–190.

Jordan V. L. 1970. Acoustical criteria for auditoriums and their relation to model techniques. *J. Acoust. Soc. Am.* Vol. **47**. pp. 408–412.

Jordan V. L. 1980. *Acoustical design of concert halls and theatres.* Applied Science Publishers Ltd.

Kleiner M. 1989. A New Way of Measuring the Lateral Energy Fraction. *Applied Acoustics* Vol. **27** (1989). pp. 321–327.

Kob M., Borländer M. 2000. Band Filters and Short Reverberation Times. *Acustica* Vol. **86** (2000). pp. 350–357.

Kovitz P. 1992. Two Maximal Length Sequence Devices for Measuring Room Acoustics Parameters. *Proc. AES 11th Int. Conf.*, May 1992.

Kutruff H. 1979. *Room Acoustics.* Second Edition. Applied Science Publishers Ltd. 309 p.

Kürer R. 1969. *Zur Gewinnung von Einzahlkriterien bei Impulsmessungen in der Raumakustik.* Acustica, Vol **21**.

Lahti T. 1990. *Analysis methods for acoustical systems based on FFT and intensity techniques.* Technical Research Centre of Finland, Publications 67. 288 p.

Lempel A. 1975. Matrix Factorization over GF(2) and Trace-Orthogonal Bases of GF(2"). *SIAM J. Comput.* Vol. **4**, No. 2, June 1975. pp. 175–186.

Lempel A., Cohn M., Eastman W. L. 1977. A Class of Balanced Binary Sequences with Optimal Autocorrelation Properties. *IEEE Trans. on Information Theory*, Vol. **IT-23**, No. 1, January 1977. pp. 38–42.

Lempel A. 1979. Hadamard and M-Sequence transforms are permutationally similar. *Applied Optics* Vol. **18**, No. 24. 15 December 1979. pp. 4064–4065.

Lundeby A., Vigran T.E., Bietz H., Vorländer M. 1995. Uncertainties of Measurements in Room Acoustics. *Acustica* Vol. **81** (1995). pp. 344–355.

The Mathworks 1998. *Matlab 5 Application Program Interface Guide*. The Mathworks, Inc. Natick, MA.

The Mathworks 1999. *Matlab 5 Function Reference*. The Mathworks, Inc. Natick, MA.

Microsoft 1997. *Microsoft Developer Network Library*. Microsoft Visual Studio™ 97 CD-ROM set.

Mommertz E., Bayer G. 1995. PC-Based High Frequency Range M-Sequence Measurements Using an Interleaved Sampling Method. *Acustica* Vol. **81** (1995). pp. 80–83.

Mommertz E., Müller S. 1995. Measuring Impulse Responses with Digitally Pre-emphasized Pseudorandom Noise Derived from Maximum-Length Sequences. *Applied Acoustics* Vol. **44** (1995). pp. 195–214.

Nielsen J. L. 1996. Maximum-length sequence measurement of room impulse responses with high level disturbances. *AES 100th conv. preprint 4267 (R-5)*. 24 p.

Nielsen J. L. 1997. Improvement of signal-to-noise ratio in long-term mls measurements with high-level nonstationary disturbances *J. Audio Eng. Soc.*, Vol. **45**, No. 12, 1997 December. pp. 1063–1066.

Oppenheim, A. V., Schafer R.W. 1975. *Digital Signal Processing*. Prentice-Hall. 585 p.

Oppenheim, A. V., Schafer R.W. 1989. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall. pp. 311–312

Rasmussen B., Hansen K. T. 1990. In-situ calculation of room acoustic parameters using a lap-top computer connected to a sound level meter. *NAM 1990 preprint*, June 11–13, Luleå 1990.

Rasmussen B., Rindel J., Henriksen H. 1991. Design and Measurement of Short Reverbation Times at Low Frequencies in Talks Studios. *J. Audio Eng. Soc.*, Vol. **39**, No. 1/2, 1991 January/February. pp. 47–57.

Rife D. D., Vanderkooy, J. 1989. Transfer-Function Measurement with Maximum-Length Sequences. *J. Audio Eng. Soc.*, Vol. **37**, No. 6, 1989 June. pp. 419–444.

Rife D. D. 1992. Modulation Transfer Function Measurement with Maximum-Length Sequences. *J. Audio Eng. Soc.*, Vol. **40**, No. 10, 1992 October. pp. 779–790.

Rife D. D. 1996. *MLSSA Reference Manual.* Version 10.0A. DRA Laboratories, 1996.

Sabine W. C. 1900. Architectural Acoustics. *Am. Arch. Building News* Vol. **68** (1900). (Reprinted in Nyman T. J. (prep.) 1922/1964. *Collected Papers on Acoustics.* Dover, New York 1964.)

Satoh F., Hidaka Y., Tachibana H. 1998. Reverberation Time Directly Obtained from Squared Impulse Response Envelope. *Proc. ICA 1998.* pp. 2755–2756.

Schroeder M. R. 1965. New Method of Measuring Reverberation Time. *J. Acoust. Soc. Am.*, Vol. **37** (1965). pp. 409–412.

Schroeder M. R. 1979. Integrated-impulse method measuring sound decay without using impulses. *J. Acoust. Soc. Am.* **66**(2), Aug. 1979. pp. 497–500.

Schroeder M. R. 1981. Modulation Transfer Functions: Definition and Measurement. *Acustica* Vol. **49** (1981). pp. 179–182.

Sutter E. E. 1991. The fast m-transform: A fast computation of cross-correlations with binary m-sequences. *SIAM J. Comput.* Vol. **20**, No. 4, August 1991. pp. 686–694.

Thiele 1953. Richtungsverteilung und Zeitvolge der Schallrückwürfe in Raümen. *Acustica* Vol. **3**. pp. 291–301. (Referenced to in [Ando 1985].)

Vanderkooy J. 1994. Aspects of MLS Measuring Systems. *J. Audio Eng. Soc.*, Vol. **42**, No. 4, 1994 April. pp. 219–231.

Vorländer M., Bietz H. 1994. Comparison of Methods for Measuring Reverberation Time. *Acustica* Vol. **80**. pp. 205–215.

Vorländer M., Kob M. 1996. Practical Aspects of MLS Measurements in Building Acoustics. *Acustica* Vol. 82 (1996) Suppl. 1. p. S103.

Vorländer M., Kob M. 1997. Practical Aspects of MLS Measurements in Building Acoustics. *Applied Acoustics* Vol. **52** (1997), No. 3/4. pp. 239–258.

MacWilliams F. J., Sloane N. J. A. 1976. Pseudo-Random Sequences and Arrays. *Proc. IEEE*, Vol. **64**, No. 12, December 1976. pp. 1715–1729.

WinMLS 2000. The WinMLS homepage.
URL: <http://www.nvo.com/winmls/door/>. 20.11.2000.

Xiang N. 1995. Evaluation of reverbation times using a nonlinear regression approach. *J. Acoust. Soc. Am.* **98**(4), 1995 October. pp. 2112–2121.

Xiang N., Genuit K. 1996. Characteristic Maximum-Length Sequences for the Interleaved Sampling Method. *Acustica* Vol. **82** (1996). pp. 905–907.

Appendix A.
MLS orders, subtypes and tap combinations

Table A.1. Implemented MLS orders, subtypes and their tap combinations.

order *n*	length *L*	subtype	taps
2	3	*a*	2, 1
3	7	*a*	3, 1
		b	3, 2
4	15	*a*	4, 1
		b	4, 3
5	31	*a*	5, 2
		b	5, 3
6	63	*a*	6, 1
		b	6, 5
7	127	*a*	7, 1
		b	7, 6
8	255	*a*	8, 6, 5, 1
		b	8, 5, 3, 2
9	511	*a*	9, 4
		b	9, 5
10	1023	*a*	10, 3
		b	10, 7
11	2047	*a*	11, 2
		b	11, 9
12	4095	*a*	12, 7, 4, 3
		b	12, 11, 8, 6
		c	12, 11, 10, 2
13	8191	*a*	13, 4, 3, 1
		b	13, 12, 10, 9
		c	13, 12, 11, 9, 6, 5, 2, 1
14	16383	*a*	14, 12, 11, 1
		b	14, 13, 8, 4
		c	14, 13, 12, 2
		d	14, 13, 12 ,10, 9, 7, 5, 3 ,1
15	32767	*a*	15, 1
		b	15, 14
		c	15, 11
		d	15, 8
		e	15, 12, 11, 8, 7, 6, 4, 2
16	65535	*a*	16, 5, 3, 2
		b	16, 15, 13, 4
		c	16, 12, 11, 10, 7, 4, 3, 2

(continued on next page)

(Table A.1. continued...)

order *n*	length *L*	subtype	taps
17	131071	*a*	17, 3
		b	17, 14
		c	17, 14, 13, 9
		d	17, 14, 11, 9, 6, 5
		e	17, 15, 13 ,11, 10, 9, 8, 4, 2, 1
18	262143	*a*	18, 7
		b	18, 11
19	524287	*a*	19, 6, 5, 1
		b	19, 18, 17, 14
20	1048575	*a*	20, 3
		b	20, 17
21	2097151	*a*	21, 2
		b	21, 19
22	4194303	*a*	22, 1
		b	22, 21
23	8388607	*a*	23, 5
		b	23, 18
24	16777215	*a*	24, 4, 3, 1
		b	24, 23, 22, 17
25	33554431	*a*	25, 3
26	67108863	*a*	26, 8, 7, 1
27	134217727	*a*	27, 8, 7, 1
28	268435455	*a*	28, 3
29	536870911	*a*	29, 2
30	1073741823	*a*	30, 16, 15, 1
31	2147483647	*a*	31, 3
32	4294967295	*a*	32, 28, 27, 1

Sequence subtypes are indexed using letters [*a*, *b*,...], corresponding to the MLS Subtype setting in the IRMA GUI. The `mls.dll` MEX function takes integers [0,1,...] for specifying the subtype parameter. The two enumeration methods are equivalent, so subtype *a* corresponds to 0, *b* to 1, etc.

The feedback taps are indexed from 1 to *n*, with *n* corresponding to the output tap of the generating register.

The tap combinations have been gathered from [MacWilliams 1976], [Vanderkooy 1994], and [Kovitz 1982].

Appendix B. IRMA hardware specifications

Table B.1. Measurement computer configuration.

part	specification	comments
cpu	333 MHz Pentium II	
memory	256 MB	DIMM
motherboard	Intel 440 LX chipset	single card computer with passive PCI backplane
hdd	10 GB EIDE	IBM
fdd	3.5", 1.44 MB	
ports	2 serial, 1 parallel, 1 USB	
network	100 Mbit Fast Ethernet	3Com
display	svga	Connected to Nokia 15" TFT flat panel display
case	19" rack, 2U	industrial case
power supply	200 W	

Table B.2. Sound card connections.

connection type	inputs	outputs	interface type	connectors
analog	2 ch	2 ch	unbal. –10/+4 dB line level	6.3 mm stereo jacks
digital	2 ch	2 ch	coaxial s-pdif (16/20 bit data)	rca
	8 ch	8 ch	optical t-dif (adat interface)	toslink
clock sync	1	1	TTL	bnc

Table B.3. Sound card specifications.[1]

Card Specifications	PCI revision 2.1 compliant, full length, installable in any MacOS-compatible or Windows 95-compatible computer equipped with full-length PCI slots
Analog Inputs	20-bit Enhanced dual bit Delta Sigma*
Analog Outputs	18-bit linear
Analog Levels	+4 dBu or –10 dBV, switchable
Frequency Response	20 Hz – 20 kHz, +0 dB, –0.6 dB
S/(THD+N) Ratio (A weighted), Input to Output	90 dB
Dynamic Range, Input to Output	94 dB
THD+N @ 1kHz (A weighted), Input to Output	0.009%
Input Impedance	1M ohm
Output impedance	50 ohm
Headroom	12 dB

Measured in accordance with FCC class A Part 15.
*The current card driver software is only capable of providing 16 bits for analog I/O.

[1] Korg Inc. 1997. *1212 I/O PCI Multichannel Interface Owner's Manual.* August 1997.

Appendix C. The IRMA irsetup structure

General

Tag	Typical value	Comment
`irsetup.description`	`'This is the default IRMA setup'`	description of current setup GUI: Files / Setup Description
`irsetup.verbose`	`1`	switch for verbose mode (commented output)
`irsetup.debug`	`0`	switch for debug mode (included for testing purposes)

File settings

Tag	Typical value	Comment
`irsetup.file.rootpath`	`'c:\irma\main\'`	root path of IRMA system, to be edited during installation (constant)
`irsetup.file.wavepath`	`[irsetup.file.rootpath 'wav\']`	path for storing stimulus and response wave files during measurement GUI: Files / Wave File Path
`irsetup.file.matfilepath`	`[irsetup.file.rootpath 'mat\']`	path for storing .mat files for filter data
`irsetup.file.setuppath`	`[irsetup.file.rootpath 'setup\']`	path for storing setups in .mat files GUI: Files / Setup File Path
`irsetup.file.textpath`	`[irsetup.file.rootpath 'text\']`	path for storing results in text table format
`irsetup.file.recplaypath`	`[irsetup.file.rootpath]`	path for recplay.exe application
`irsetup.file.setupfile`	`'default'`	name of current setup file GUI: Files / Setup Filename
`irsetup.file.resultfile`	`'results.txt'`	name of result file to write into
`irsetup.file.playfile`	`'stimulus.wav'`	name of wave file for storing the stimulus before measurement GUI: Files / Stimulus Filename
`irsetup.file.recfile`	`'response.wav'`	name of wave file for storing the raw response during measurement GUI: Files / Response Filename
`irsetup.file.recplayinifile`	`'recplay.ini'`	name of .ini setup file for recplay.exe
`irsetup.file.appendresultfile`	`0`	set to append into existing result text file

Stimulus settings

Tag	Typical value	Comment
irsetup.stimulus.typelist	{'MLS', 'Impulse', 'Convolution'};	list of available stimuli: maximum length sequence / periodic unit impulse / arbitrary stimulus
irsetup.stimulus.type	'MLS'	selected stimulus type GUI: Stimulus / Stimulus Type
irsetup.stimulus.cycles	2	repeat sequence n times, minimum = 2 GUI: Stimulus / Repeat Sequence
irsetup.stimulus.predelaylist	[0 0 0 0 0 0]	used to compensate for "negative" loopback delay, if acquisition begins before output: inserts n samples of silence at start of stimulus (output device specific)
irsetup.stimulus.amplitudelist	[-6 -6 -6 -6 -6 -6]	stimulus levels in dB FS (output device specific) GUI: Channels / Output Levels
irsetup.stimulus.impulse.length	1000	length of unit impulse period in samples GUI: Stimulus / Cycle Length
irsetup.stimulus.mls.orderlist	2:24	available orders for MLS (constant)
irsetup.stimulus.mls.order	16	selected MLS order GUI: Stimulus / MLS Order
irsetup.stimulus.mls.subtypenamelist	{'A' 'B' 'C' 'D' 'E'}	labels for subtypes (constant)
irsetup.stimulus.mls.subtype	1	choice of tap placement, selects subtypenamelist(n) GUI: Stimulus / MLS Subtype
irsetup.stimulus.mls.subtypelist	[0 1 2 2 2 2 2 2 2 2 2 3 3 4 5 3 5 2 2 2 2 2 2 2 1 1 1 1 1 1 1]	number of available subtypes for each mls order (constant)
irsetup.stimulus.custom	[]	name of global variable containing custom stimulus for convolution GUI: Stimulus / Custom Stimulus
irsetup.stimulus.out.count	1	number of output devices in use
irsetup.stimulus.out.idlist	[4]	id's of output devices in use GUI: Channels / Active Outputs
irsetup.stimulus.out.namelist	{'optical 1-2', 'optical 3-4', 'optical 5-6', 'optical 7-8', 'analog line input', 'S/PDIF input'}	labels describing system outputs GUI: Devices / Device Output Label

Acquisition settings

Tag	Typical value	Comment
irsetup.acquisition. sampleratelist	[11025 22050 44100 48000]	possible sample rates (constant)
irsetup.acquisition.samplerate	44100	selected sample rate GUI: Stimulus / Sample Rate
irsetup.acquisition. enablepreaverage	1	set to enable preaveraging GUI: Acquisition / Enable Pre-averaging
irsetup.acquisition. excludefirst	1	set to exclude first preavg cycle GUI: Acquisition / Exclude First Cycle
irsetup.acquisition. excludelast	0	set to exclude last preavg cycle GUI: Acquisition / Exclude Last Cycle
irsetup.acquisition. preaveragelist	1:irsetup.stimulus.cycles - irsetup.acquisition. excludefirst - irsetup.acquisition. excludelast	possible values for number of preaveraged cycles
irsetup.acquisition.preaverage	max(irsetup.acquisition. preaveragelist)	number of cycles to be preaveraged
irsetup.acquisition. offlinemode	0	set for "offline" acquisition: skips recplay and loads raw responses directly from the specified response.wav file GUI: Files / Work Offline
irsetup.acquisition.truncateir	0	set to truncate impulse response GUI: Acquisition / Truncate Response Length
irsetup.acquisition. truncatelength	1024	truncated response length (samples) GUI: Acquisition / Truncate Response Length to ... samples
irsetup.acquisition.in.count	1	number of input devices in use
irsetup.acquisition.in.idlist	[1]	id's of input devices in use GUI: Channels / Active Inputs
irsetup.acquisition.in. namelist	{ 'A/D ch. 1-2', 'A/D ch. 3-4', 'A/D ch. 5-6', 'A/D ch. 7-8', 'Line Input', 'spdif'}	input labels GUI: Devices / Device Input Label
irsetup.acquisition.compensate	0	set to compensate for loopback response GUI: Acquisition/Compensate IR
irsetup.acquisition.loopbackir	''	name of workspace variable for loopback response GUI: Acquisition / Loopback Response in Workspace Variable
irsetup.acquisition.ir	'ir'	name of global variable for IR GUI: Acquisition / Store Measured Response in Workspace Variable

System sound device settings

Tag	Typical value	Comment
irsetup.device.out.count	6	number of output devices in system (constant)
irsetup.device.out.idlist	[0 1 2 3 4 5]	list of device id's (constant)
irsetup.device.out.namelist	{'Korg 1212 I/O Adat 1-2', 'Korg 1212 I/O Adat 3-4', 'Korg 1212 I/O Adat 5-6', 'Korg 1212 I/O Adat 7-8', 'Korg 1212 I/O Analog', 'Korg 1212 I/O S/PDIF'}	device names (constant)
irsetup.device.in.count	6	number of input devices in system (constant)
irsetup.device.in.idlist	[0 1 2 3 4 5]	list of device id's (constant)
irsetup.device.in.namelist	{'Korg 1212 I/O Adat 1-2', 'Korg 1212 I/O Adat 3-4', 'Korg 1212 I/O Adat 5-6', 'Korg 1212 I/O Adat 7-8', 'Korg 1212 I/O Analog', 'Korg 1212 I/O S/PDIF'}	device names (constant)
irsetup.device.in.sampleoffset	[102 102 102 102 102 102]	input loopback delays in samples (device specific) GUI: Devices / Delay

Analysis settings

Tag	Typical value	Comment
irsetup.analysis.results.snr	1	calculate signal to noise ratios
irsetup.analysis.results.rt20	1	calculate reverberation by rt20
irsetup.analysis.results.rt30	1	calculate reverberation by rt30
irsetup.analysis.results.rtuser	1	calculate rt with user parameters
irsetup.analysis.results.edt	1	calculate early decay time
irsetup.analysis.results.c50	1	calculate c50
irsetup.analysis.results.c80	1	calculate c80
irsetup.analysis.results.d50	1	calculate d50
irsetup.analysis.results.d80	0	calculate d80
irsetup.analysis.results.euser	0	calculate energy with user parameters
irsetup.analysis.results.ts	1	calculate center time
irsetup.analysis.results.level	0	calculate level
irsetup.analysis.results.lef	0	calculate lateral energy fraction
irsetup.analysis.results.iacc	0	calculate interaural cross correlation
irsetup.analysis.rtuserlimits	[-10 -40]	rt_user dB limits
irsetup.analysis.euserlimits	[0 80 0 Inf]	e_user dB limits

Filter settings

Tag	Typical value	Comment
irsetup.filter.fcoctlist	[125 250 500 1000 2000 4000 8000]	list of octave center frequencies (constant)
irsetup.filter.fcthirdlist	[100 125 160 200 250 315 400 500 630 800 1000 1250 1600 2000 2500 3150 4000 5000 6300 8000 10000 12500]	list of third oct center frequencies (constant)
irsetup.filter.fcoct1	1	octave start frequency, fclist(n) GUI: Filters / Range
irsetup.filter.fcoct2	7	octave end frequency, fclist(n) GUI: Filters / Range
irsetup.filter.fcthird1	1	third oct end frequency, fclist(n) GUI: Filters / Range
irsetup.filter.fcthird2	18	third oct end frequency, fclist(n) GUI: Filters / Range
irsetup.filter.filtir	'filtir'	variable name for storing filtered ir's GUI: Filters / Store Filtered Responses in Workspace Variable
irsetup.fcthirdoffset	20	ISO third oct index for fcthirdlist(1) (constant)
irsetup.filter.type	3	1=octaves, 3=thirds GUI: Filters / Bandpass Filtering
irsetup.filter.datachannel	1	filter data in channel n GUI: Filters / Filter Data in Channel
irsetup.filter.fcnames	[]	list of center frequencies for the current filter selection
irsetup.filter.timereverse	1	set for time-reversed filtering